SYSTEM, METHOD AND ARTICLE OF MANUFACTURE FOR UPDATING A SWITCHING TABLE IN A SWITCH FABRIC CHIPSET SYSTEM

Inventors: You-Sung CHANG, et al.

Attorney Docket No.: 56030.00027

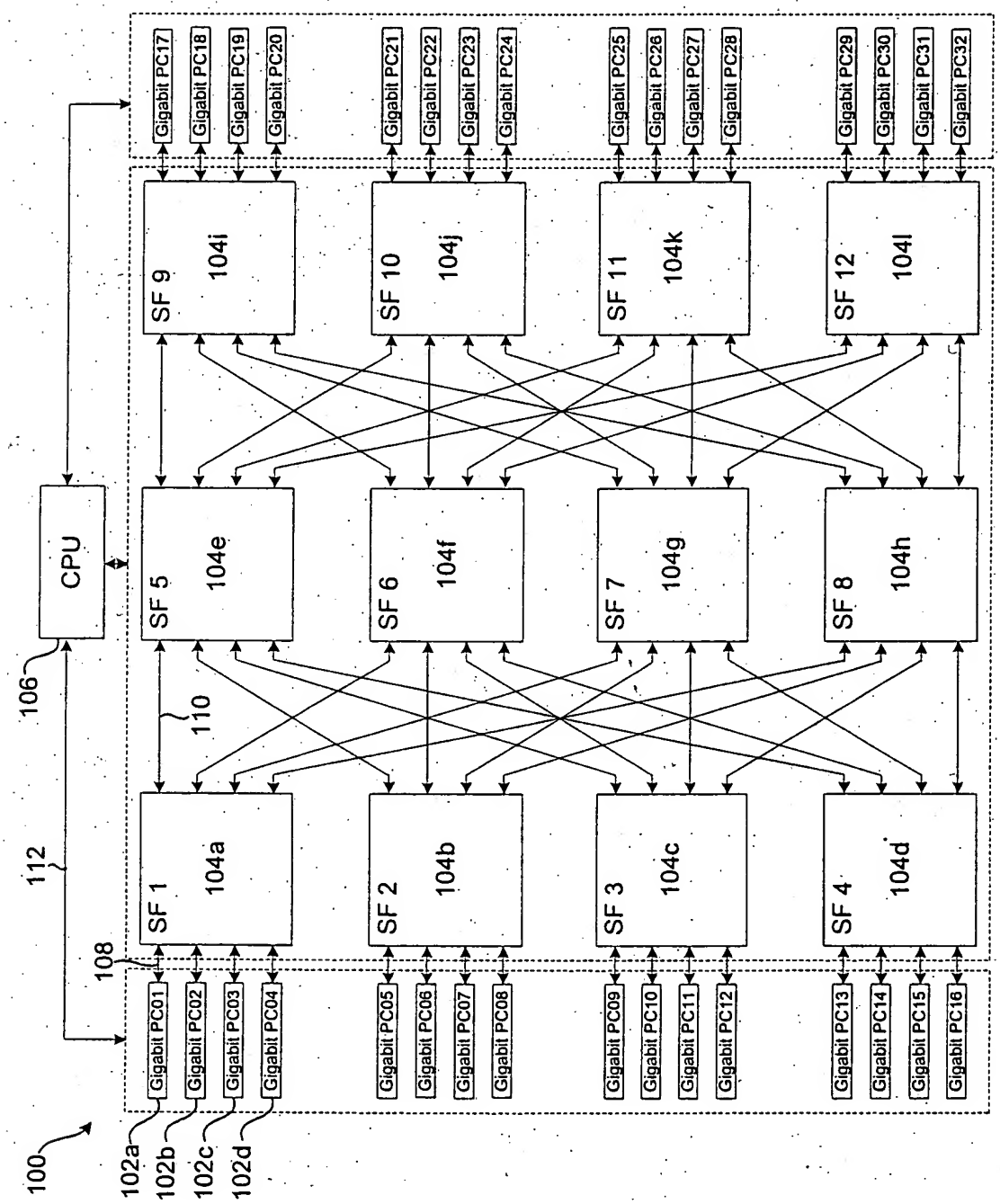

FIG. 1

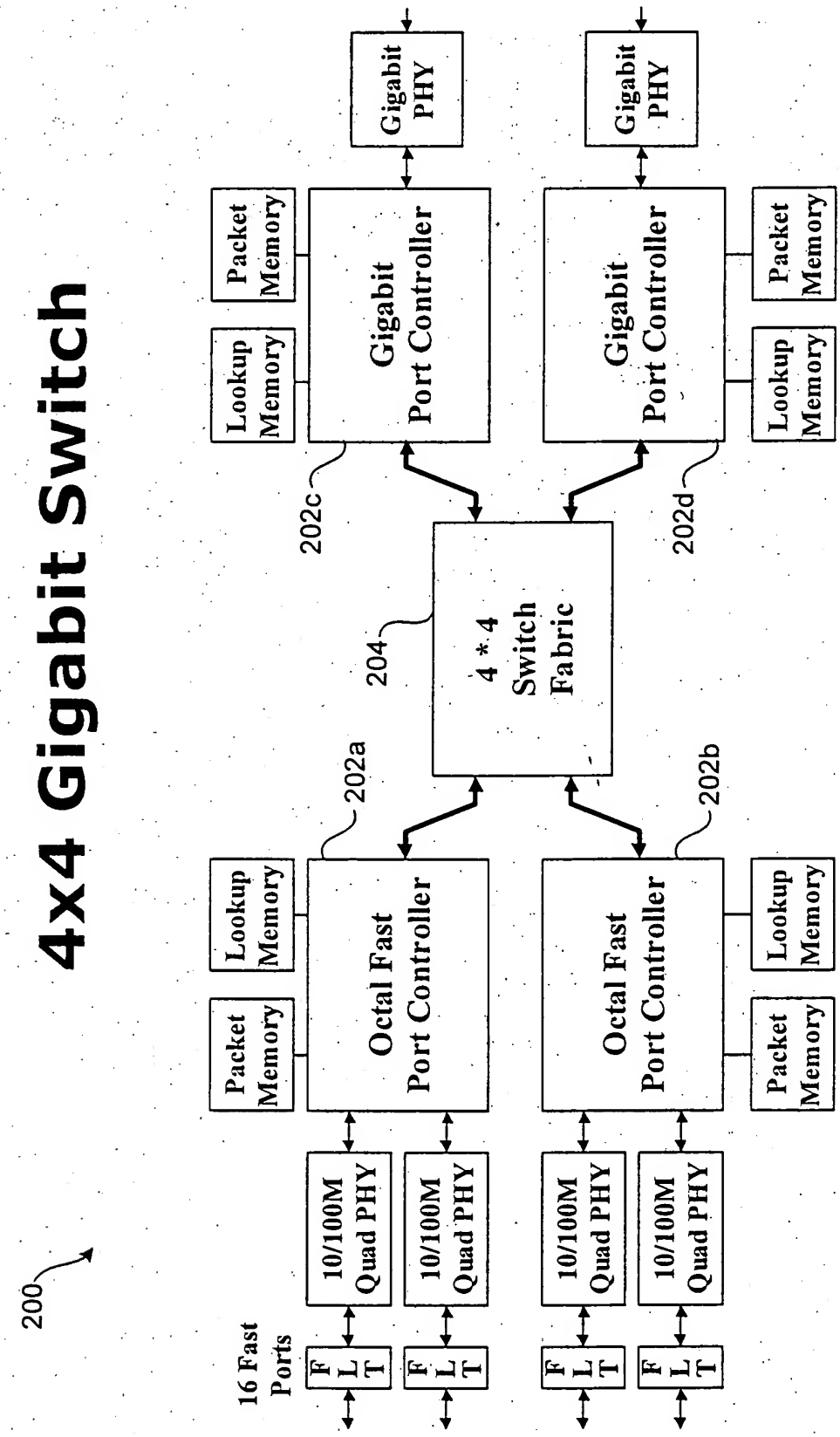
4x4 Gigabit Switch
200
16 Fast Ports
F L T
10/100M Quad PHY
Octal Fast Port Controller
Packet Memory
Lookup Memory
202a
202b
4 * 4 Switch Fabric
204
Gigabit Port Controller
Gigabit PHY
202c
202d

FIG. 2

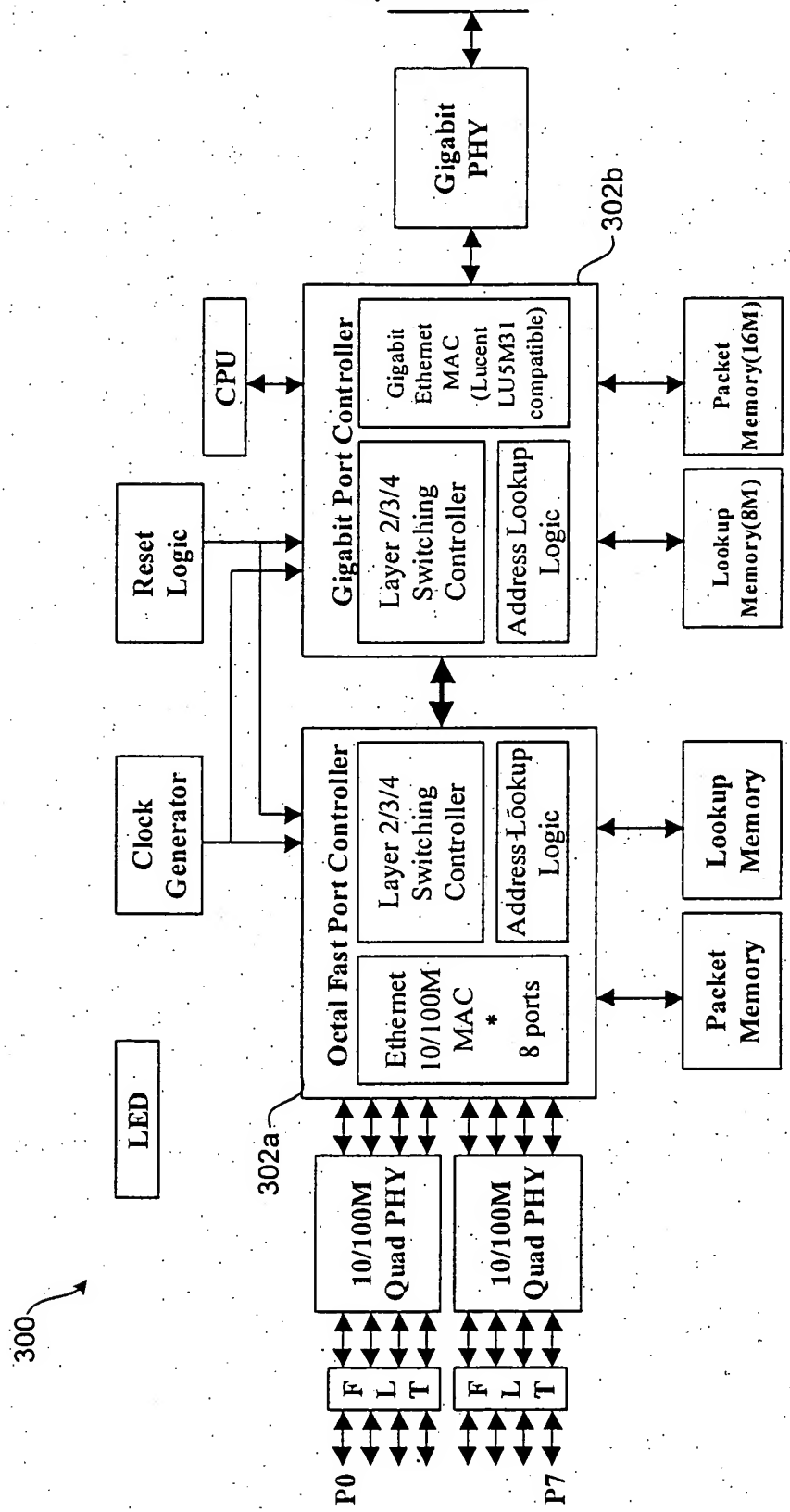
Gigabit Up-linking
300
LED
Clock Generator
Reset Logic
CPU
Gigabit PHY
302b
Gigabit Port Controller
Gigabit Ethernet MAC (Lucent LU5M31 compatible)
Layer 2/3/4 Switching Controller
Address Lookup Logic
Packet Memory(16M)
Lookup Memory(8M)
Octal Fast Port Controller
Layer 2/3/4 Switching Controller
Address Lookup Logic
Ethernet 10/100M MAC * 8 ports
Lookup Memory
Packet Memory
302a
10/100M Quad PHY
10/100M Quad PHY
F L T
F L T
P0
P7

FIG. 3

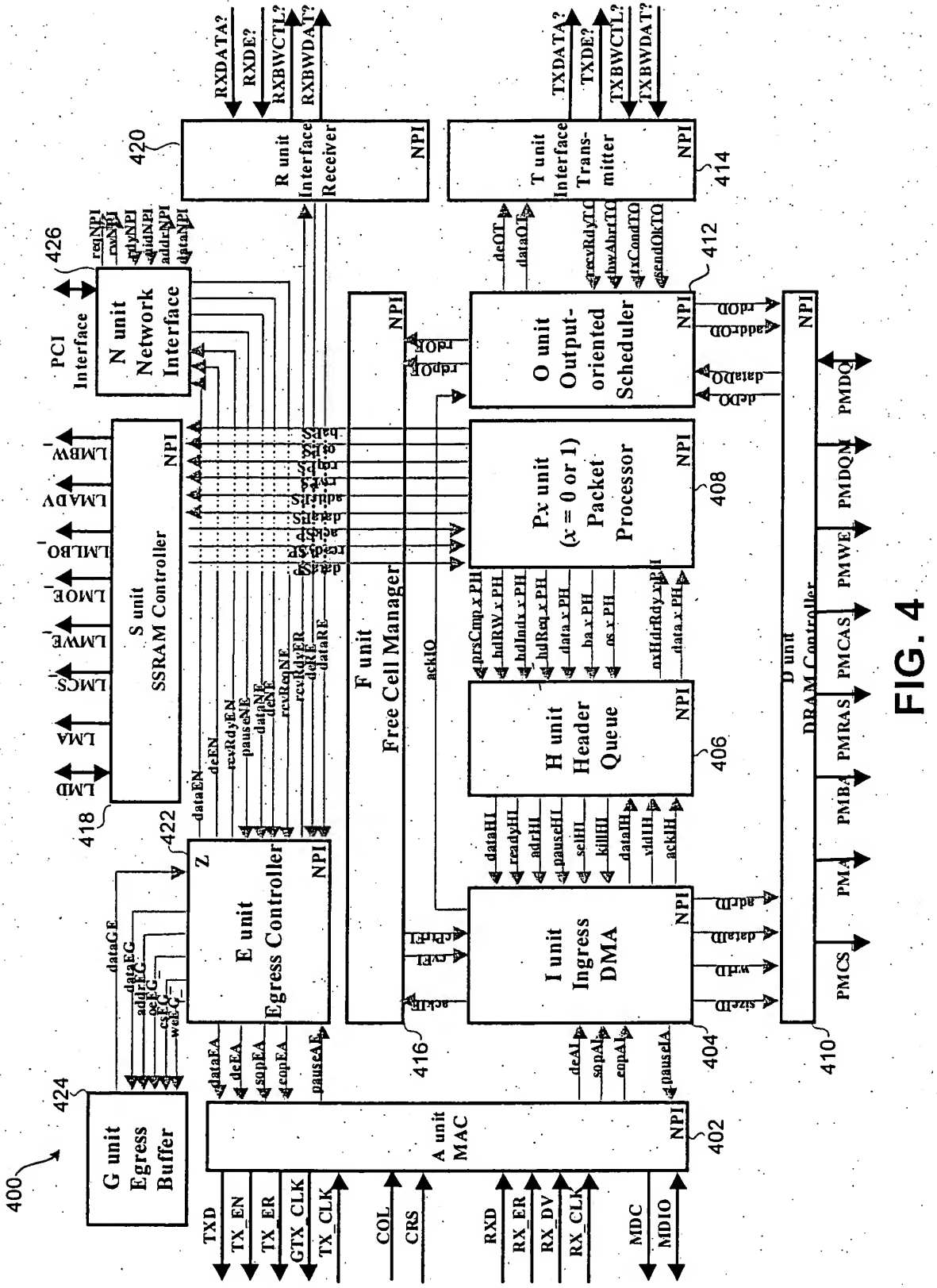
400
G unit Egress Buffer
424
A unit MAC
402
E unit Egress Controller
422
I unit Ingress DMA
404
H unit Header Queue
406
Px unit (x = 0 or 1) Packet Processor
408
F unit Free Cell Manager
416
O unit Output-oriented Scheduler
412
T unit Interface Trans-mitter
414
R unit Interface Receiver
420
N unit Network Interface
426
PCI Interface
S unit SSRAM Controller
418
D unit DRAM Controller
410
NPI
RXDATA?
RXDE?
RXBWCTL?
RXBWDAT?
TXDATA?
TXDE?
TXBWCTL?
TXBWDAT?
TXD
TX_EN
TX_ER
GTX_CLK
TX_CLK
COL
CRS
RXD
RX_ER
RX_DV
RX_CLK
MDC
MDIO
LMD
LMA
LMCS
LMWE
LMOE
LMLBO
LMADV
LMBW
PMDQ
PMDQM
PMWE
PMCAS
PMRAS
PMBA
PMA
PMCS

FIG. 4

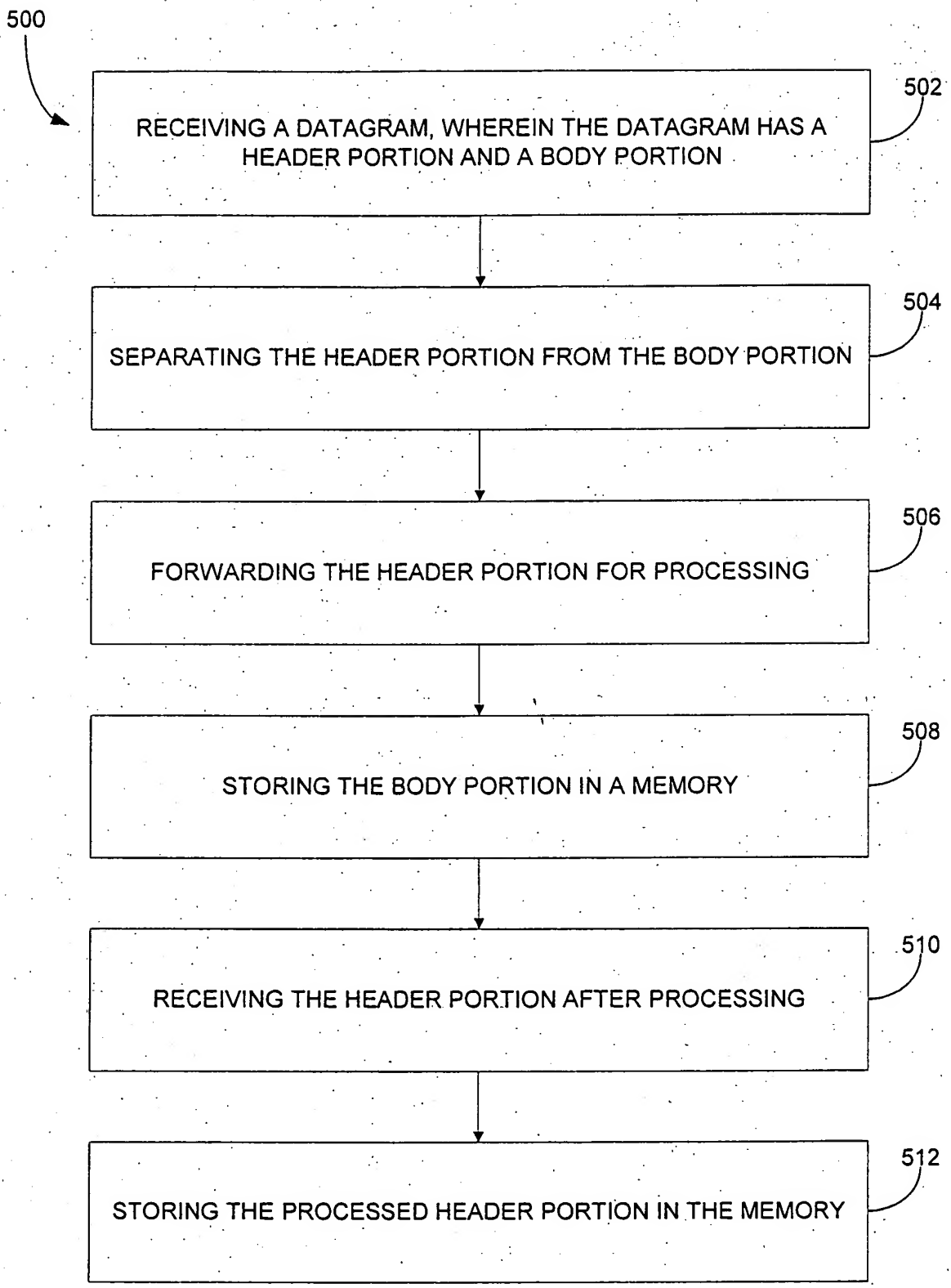
500
RECEIVING A DATAGRAM, WHEREIN THE DATAGRAM HAS A HEADER PORTION AND A BODY PORTION
502
SEPARATING THE HEADER PORTION FROM THE BODY PORTION
504
FORWARDING THE HEADER PORTION FOR PROCESSING
506
STORING THE BODY PORTION IN A MEMORY
508
RECEIVING THE HEADER PORTION AFTER PROCESSING
510
STORING THE PROCESSED HEADER PORTION IN THE MEMORY
512

FIG. 5

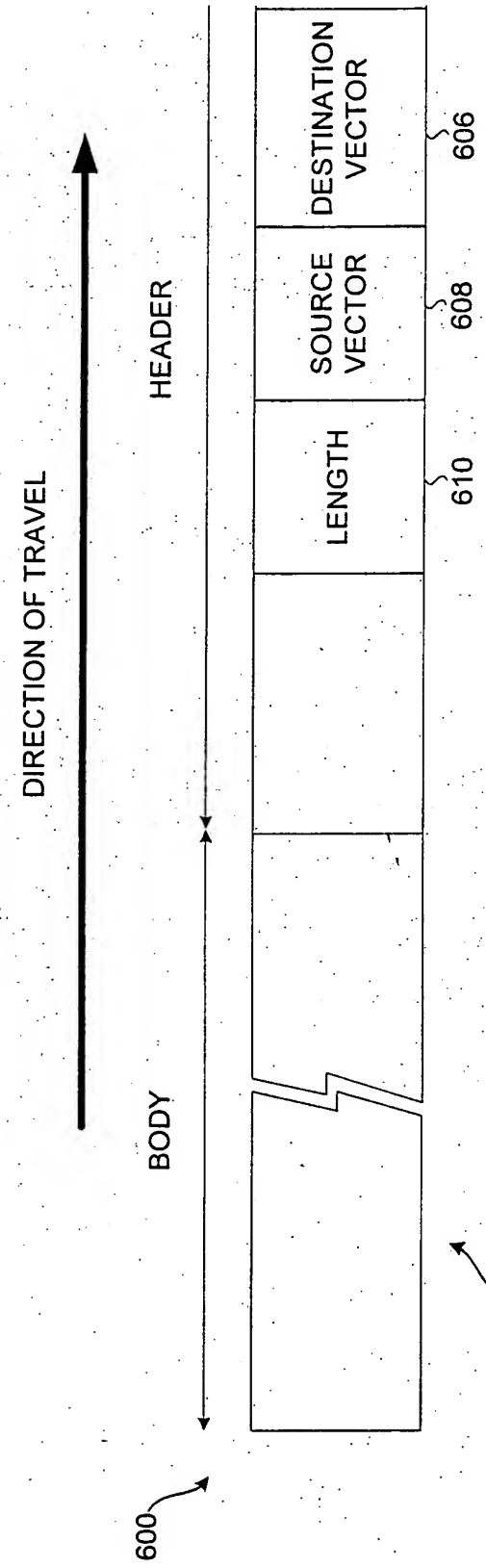
DIRECTION OF TRAVEL
HEADER
BODY
DESTINATION VECTOR
SOURCE VECTOR
LENGTH
606
608
610
602
600

FIG. 6

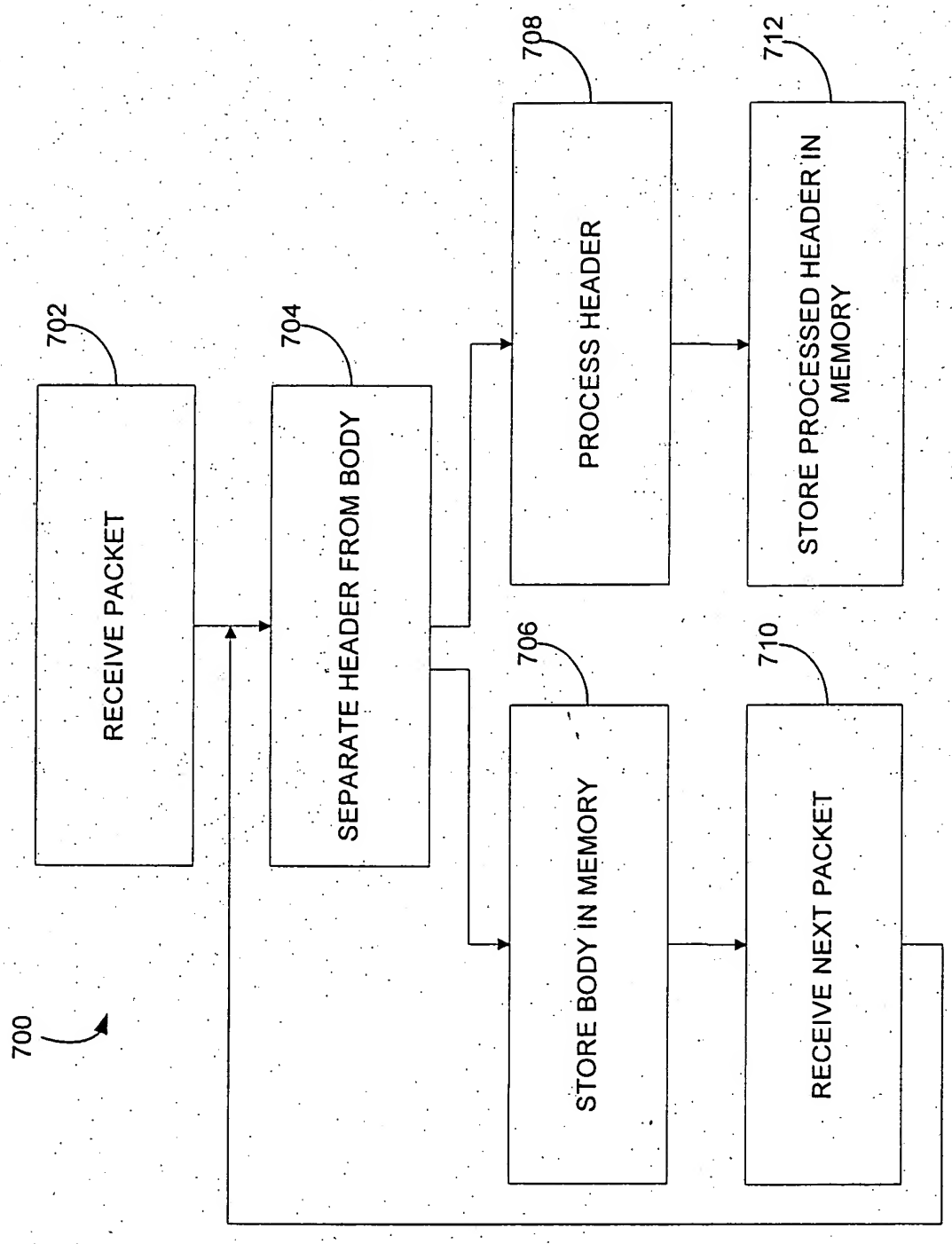
700
RECEIVE PACKET
702
SEPARATE HEADER FROM BODY
704
PROCESS HEADER
708
STORE PROCESSED HEADER IN MEMORY
712
STORE BODY IN MEMORY
706
RECEIVE NEXT PACKET
710

FIG. 7

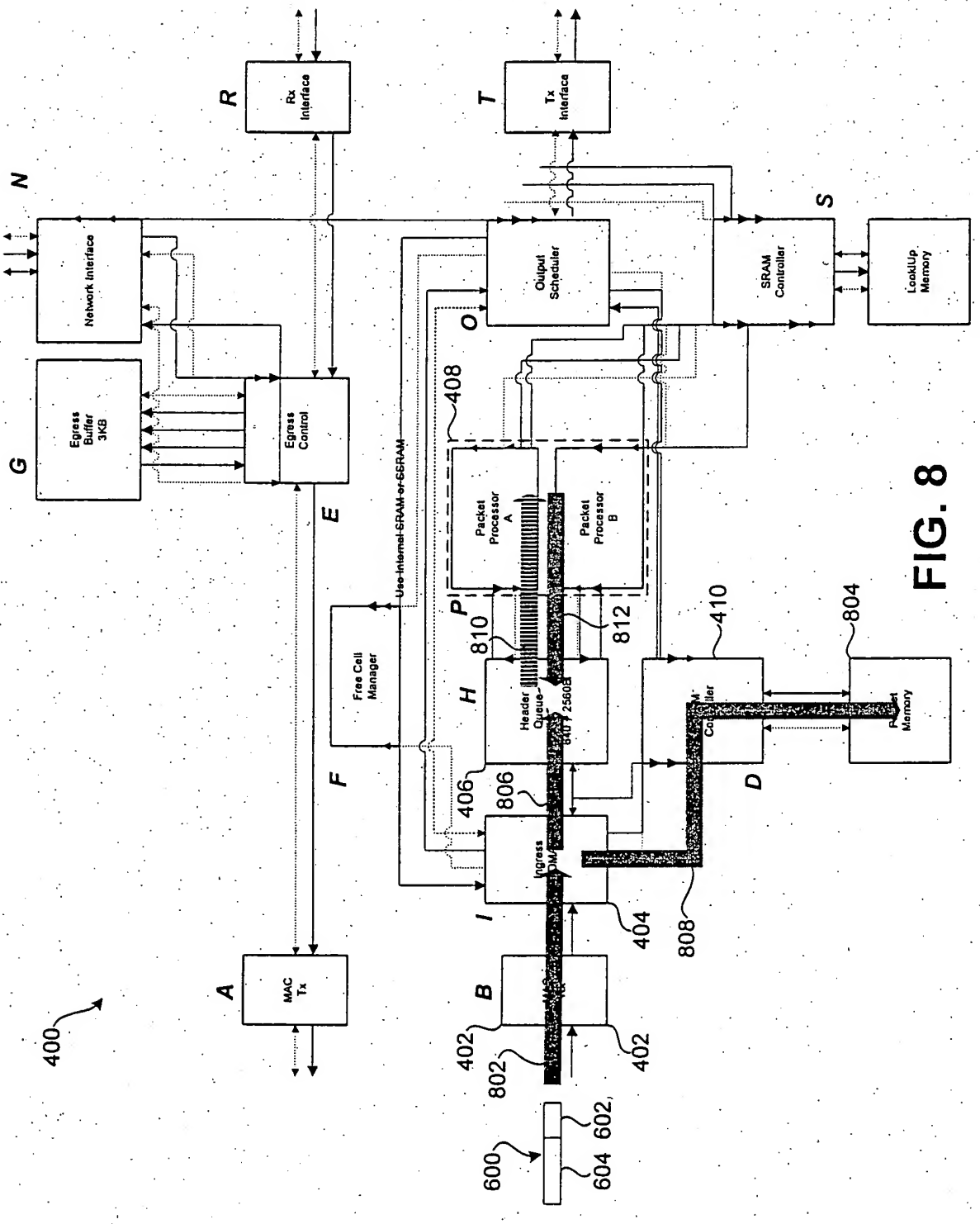
R
Rx Interface
T
Tx Interface
N
Network Interface
S
SRAM Controller
LookUp Memory
O
Output Scheduler
408
G
Egress Buffer 3KB
Egress Control
E
Packet Processor A
Packet Processor B
P
Use Internal SRAM or SSRAM
Free Cell Manager
F
H
Header Queue
810
812
410
804
Packet Memory
D
406
806
Ingress DMA
I
404
808
A
MAC Tx
B
402
802
602
600
604
400

FIG. 8

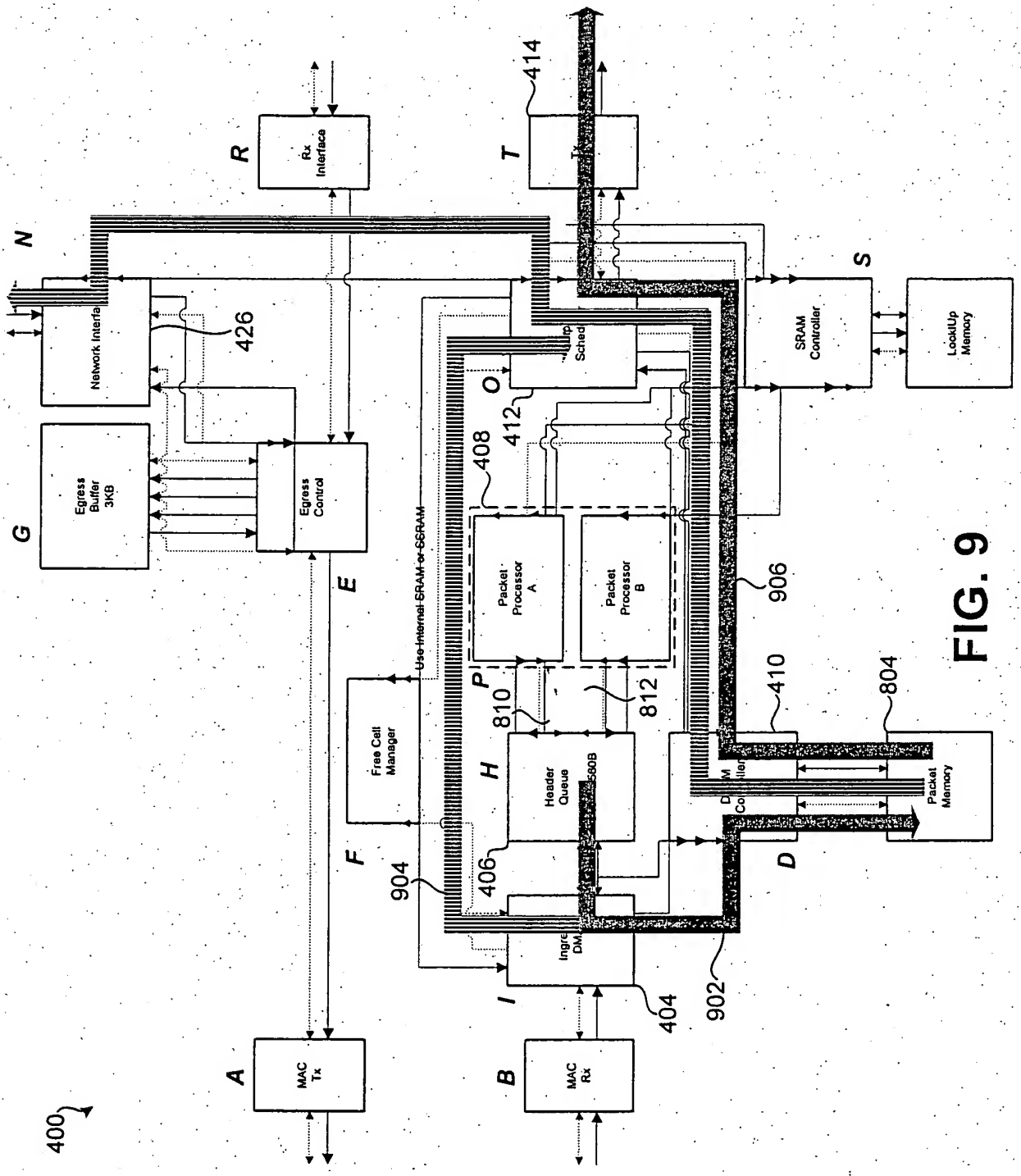
400
A
MAC Tx
B
MAC Rx
I
404
902
F
904
406
H
Header Queue
560B
Free Cell Manager
P
810
812
Packet Processor A
Packet Processor B
906
408
412
O
Egress Buffer 3KB
Egress Control
G
E
Network Interfa
426
N
R
Rx Interface
T
414
SRAM Controller
LookUp Memory
S
410
804
Packet Memory
D

FIG. 9

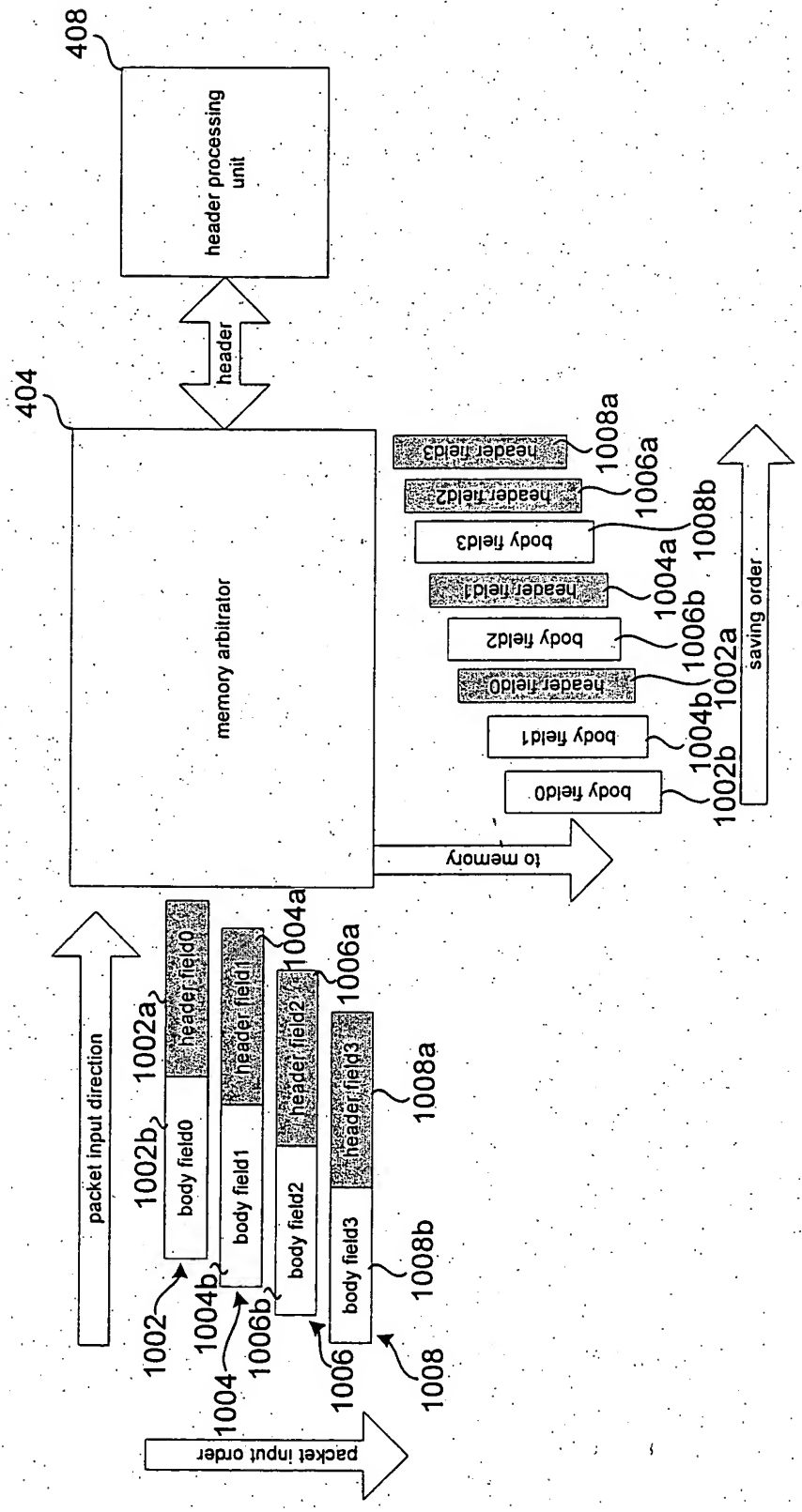
408
header processing unit
header
404
memory arbitrator
packet input direction
packet input order
to memory
saving order
header field0
header field1
header field2
header field3
body field0
body field1
body field2
body field3
1002
1002a
1002b
1004
1004a
1004b
1006
1006a
1006b
1008
1008a
1008b

FIG. 10

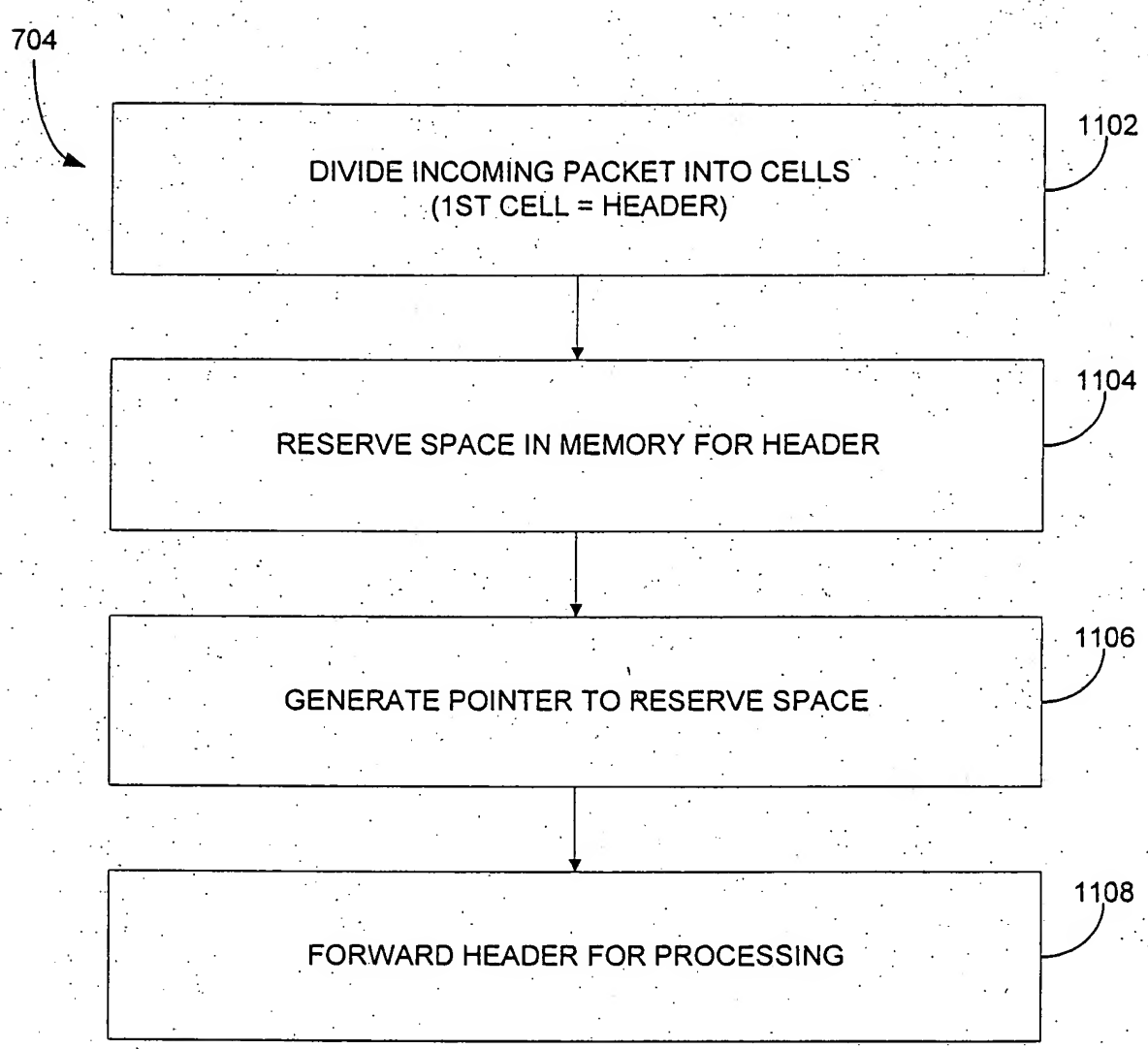
704
DIVIDE INCOMING PACKET INTO CELLS
(1ST CELL = HEADER)
1102
RESERVE SPACE IN MEMORY FOR HEADER
1104
GENERATE POINTER TO RESERVE SPACE
1106
FORWARD HEADER FOR PROCESSING
1108

FIG. 11

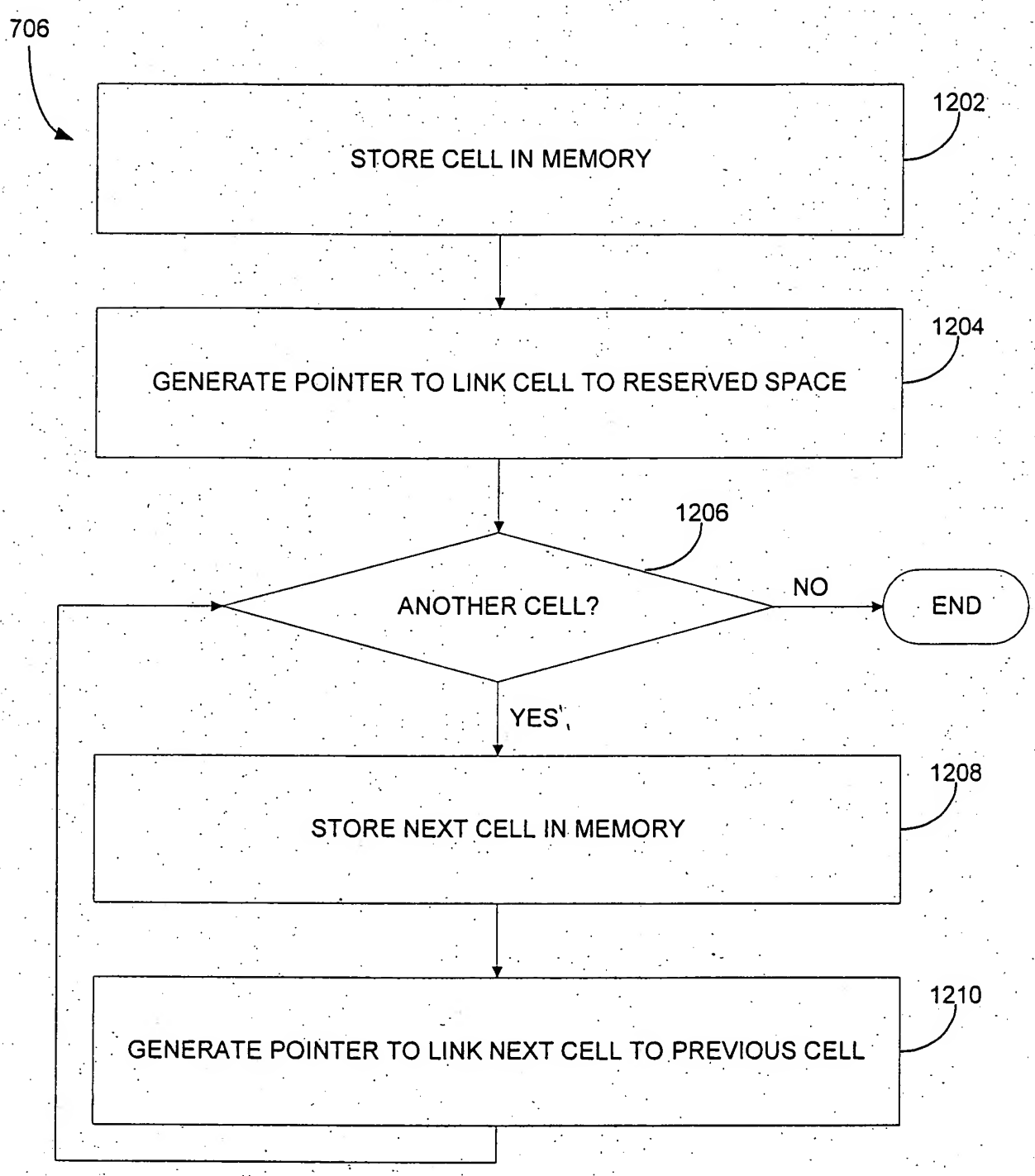
706
STORE CELL IN MEMORY
1202
GENERATE POINTER TO LINK CELL TO RESERVED SPACE
1204
1206
ANOTHER CELL?
NO
END
YES
STORE NEXT CELL IN MEMORY
1208
GENERATE POINTER TO LINK NEXT CELL TO PREVIOUS CELL
1210

FIG. 12

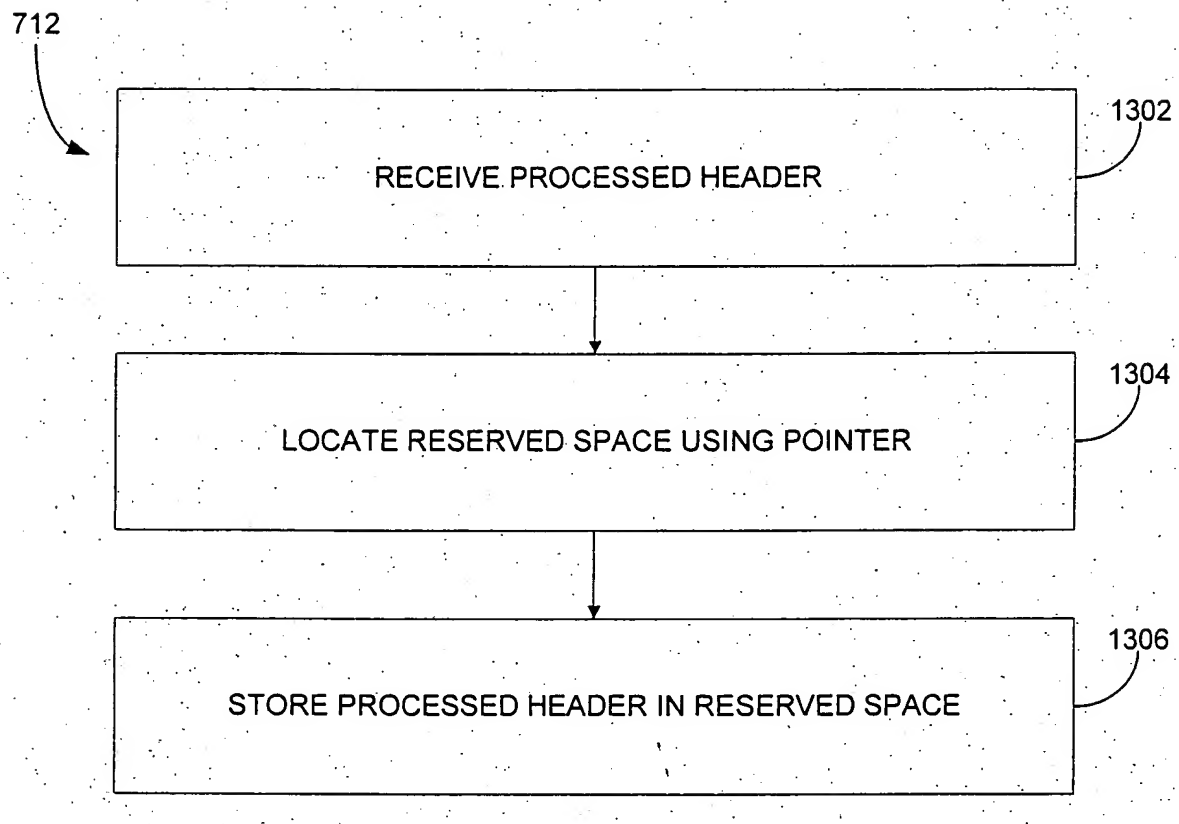
712
RECEIVE PROCESSED HEADER
1302
LOCATE RESERVED SPACE USING POINTER
1304
STORE PROCESSED HEADER IN RESERVED SPACE
1306

FIG. 13

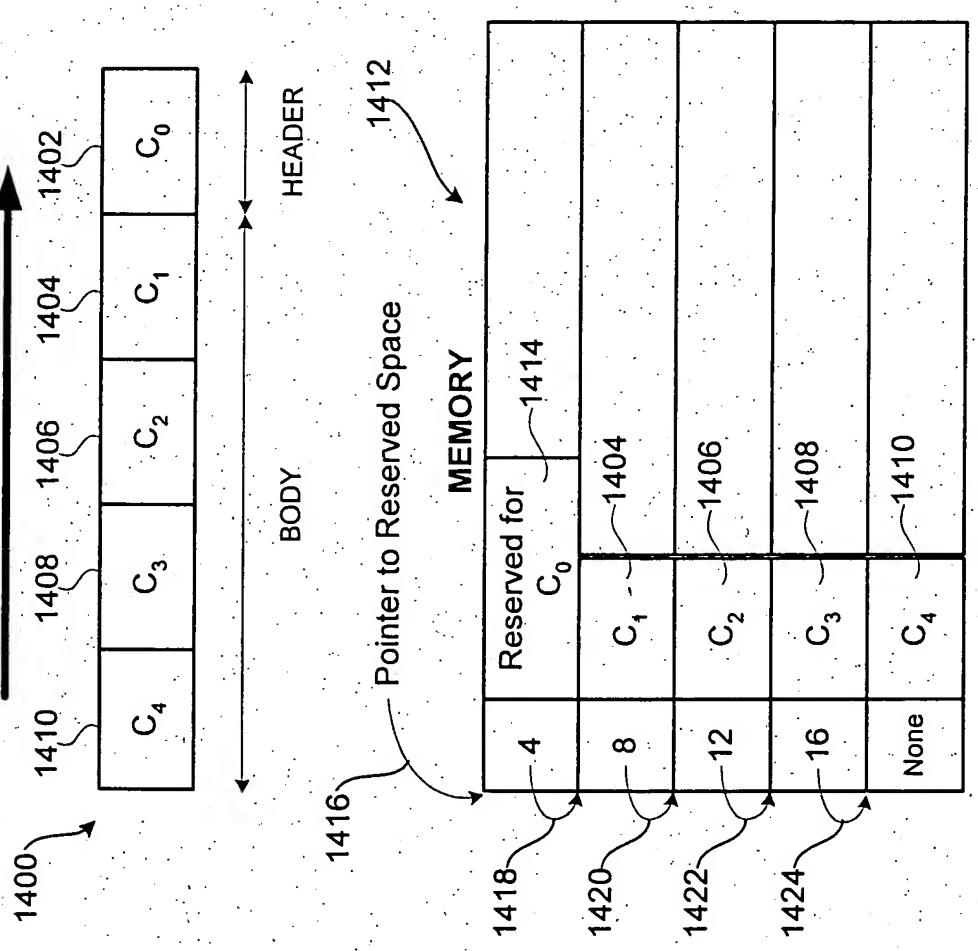
1400
1402
1404
1406
1408
1410
C_0
C_1
C_2
C_3
C_4
HEADER
BODY
1412
MEMORY
1414
Reserved for C_0
1404
1406
1408
1410
C_1
C_2
C_3
C_4
Pointer to Reserved Space
1416
4
8
12
16
None
1418
1420
1422
1424

FIG. 14

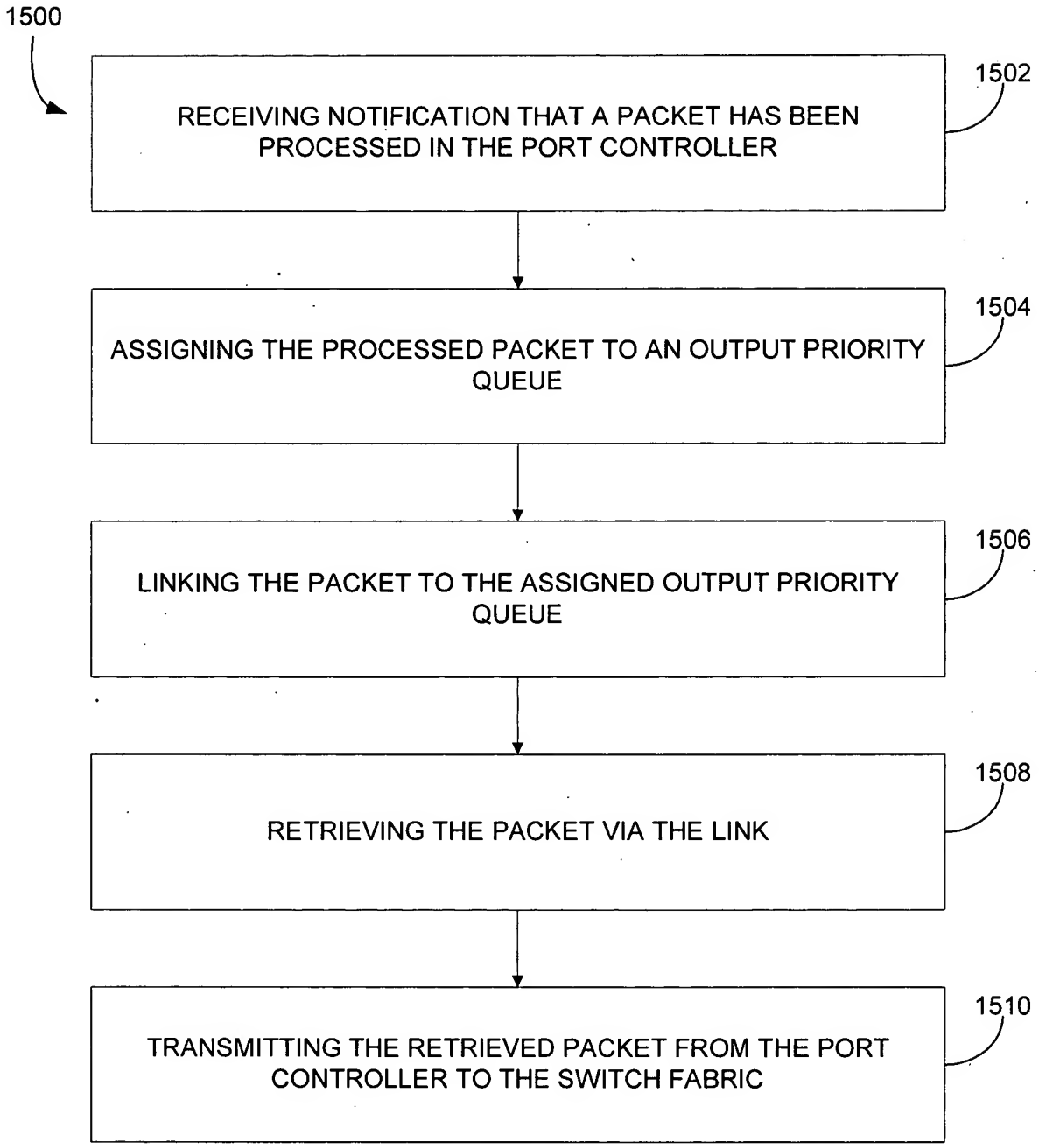
1500
RECEIVING NOTIFICATION THAT A PACKET HAS BEEN PROCESSED IN THE PORT CONTROLLER
1502
ASSIGNING THE PROCESSED PACKET TO AN OUTPUT PRIORITY QUEUE
1504
LINKING THE PACKET TO THE ASSIGNED OUTPUT PRIORITY QUEUE
1506
RETRIEVING THE PACKET VIA THE LINK
1508
TRANSMITTING THE RETRIEVED PACKET FROM THE PORT CONTROLLER TO THE SWITCH FABRIC
1510

FIG. 15

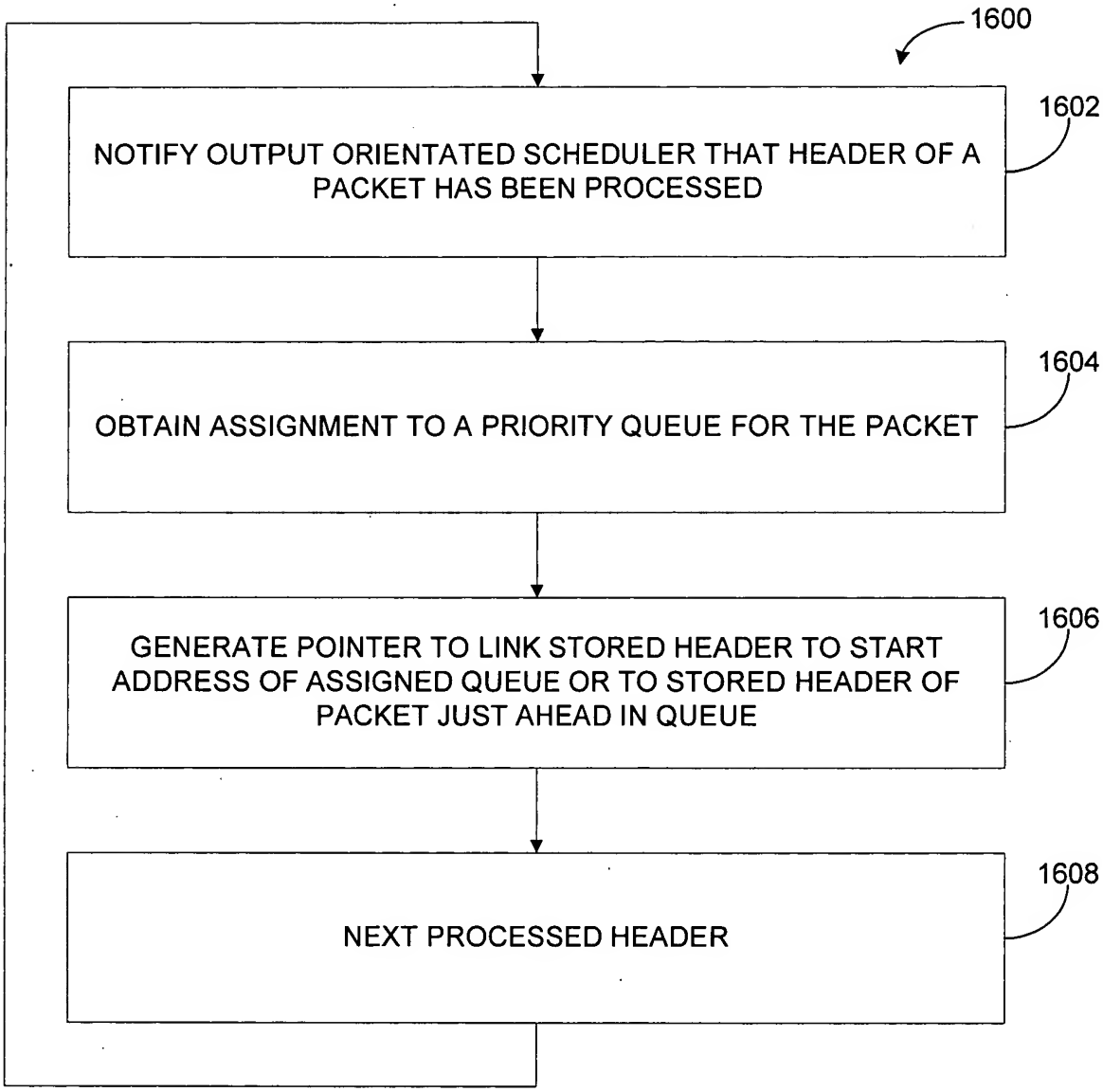
1600
NOTIFY OUTPUT ORIENTATED SCHEDULER THAT HEADER OF A PACKET HAS BEEN PROCESSED
1602
OBTAIN ASSIGNMENT TO A PRIORITY QUEUE FOR THE PACKET
1604
GENERATE POINTER TO LINK STORED HEADER TO START ADDRESS OF ASSIGNED QUEUE OR TO STORED HEADER OF PACKET JUST AHEAD IN QUEUE
1606
NEXT PROCESSED HEADER
1608

FIG. 16

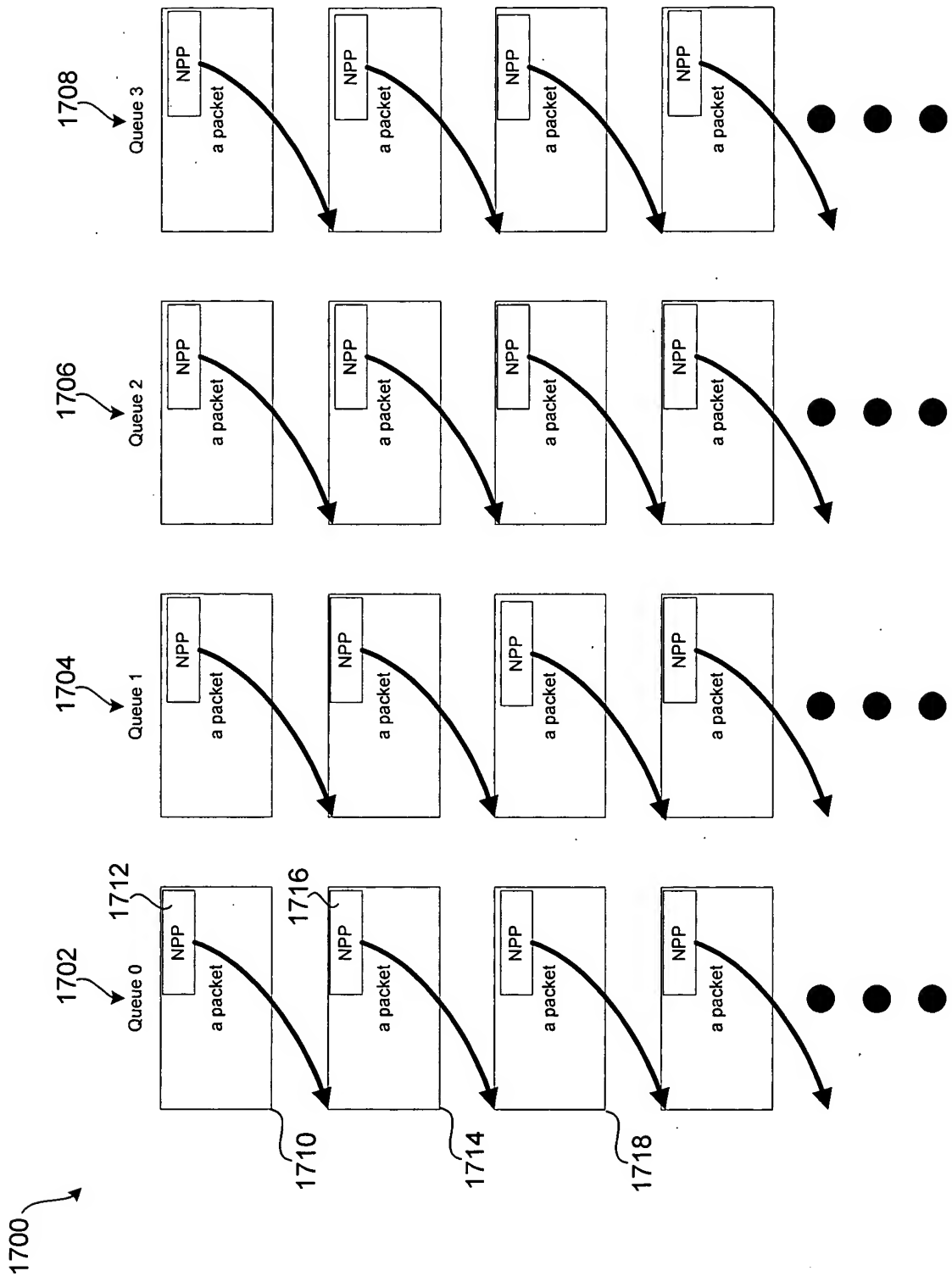
1700
1702
Queue 0
1704
Queue 1
1706
Queue 2
1708
Queue 3
1710
1712
1714
1716
1718
NPP
a packet

FIG. 17

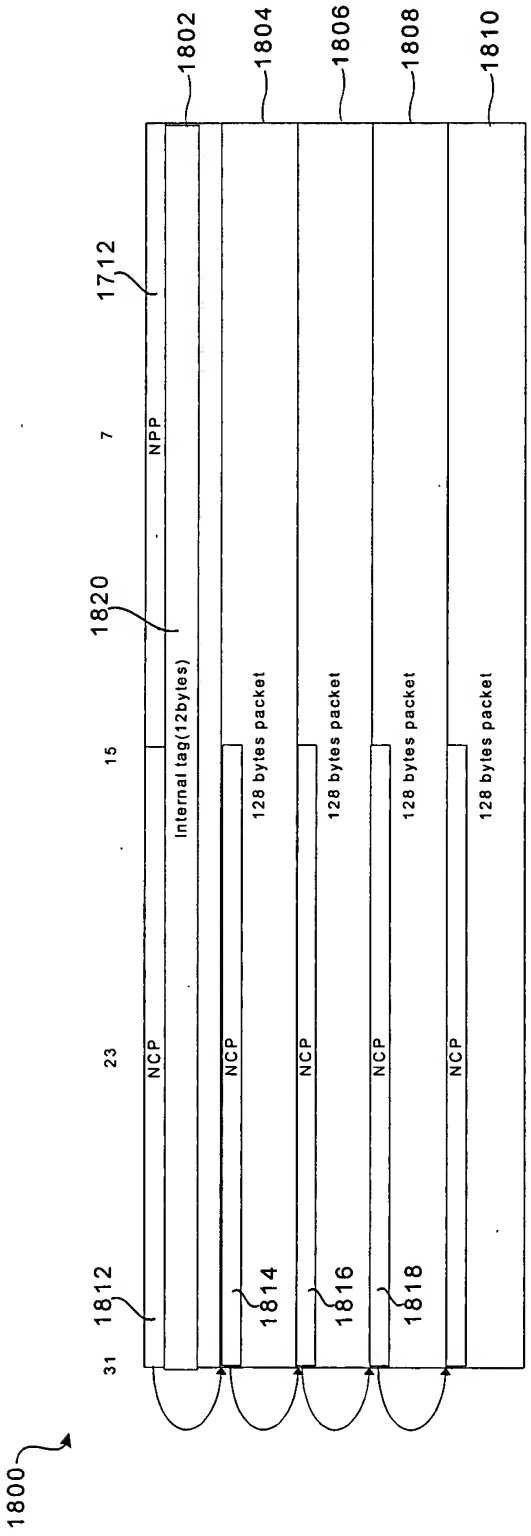
1802
1804
1806
1808
1810
1712
1820
1812
1814
1816
1818
1800
0
7
15
23
31
NPP
NCP
Internal tag(12bytes)
128 bytes packet

FIG. 18

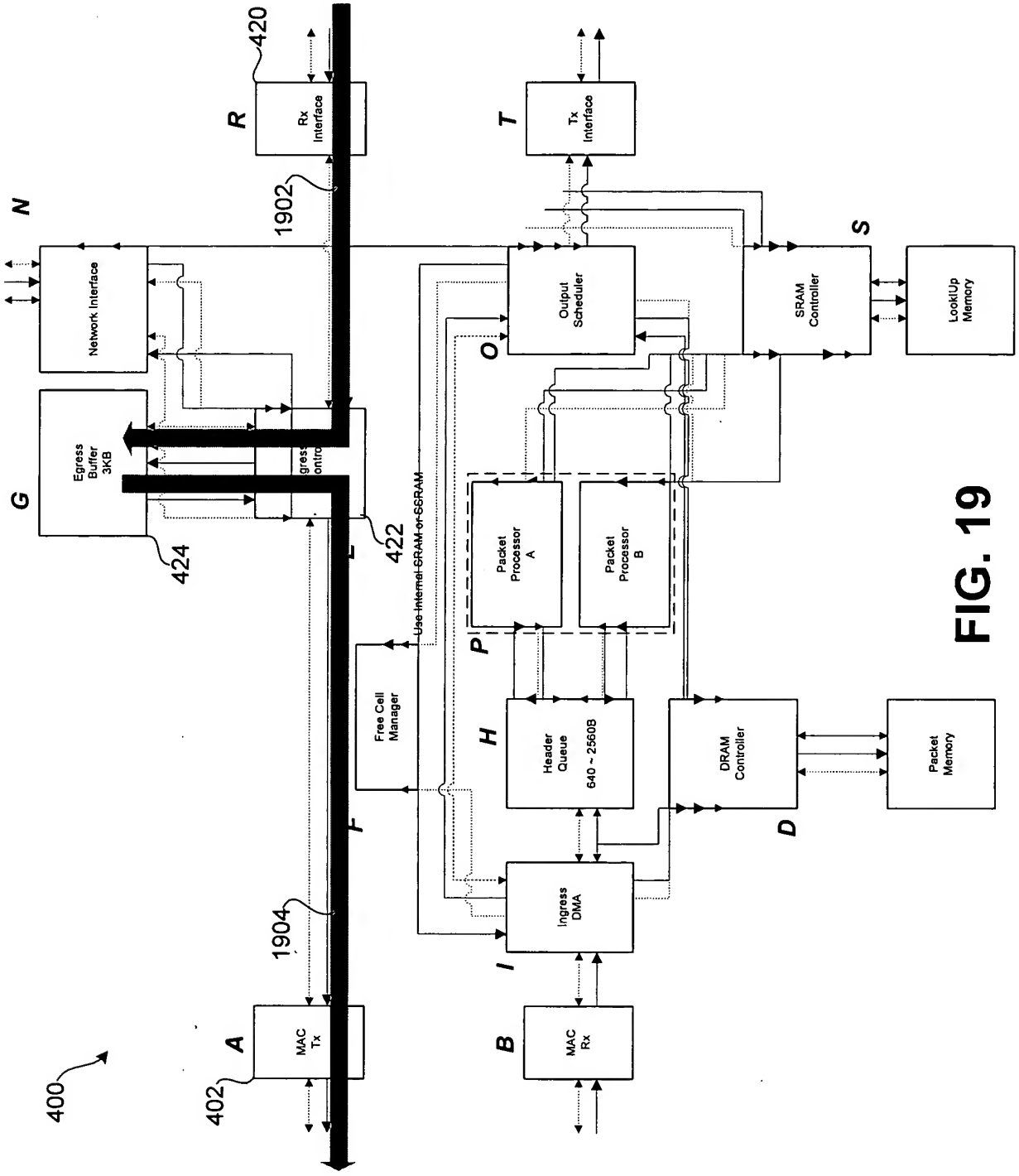
420
R
Rx Interface
T
Tx Interface
1902
N
Network Interface
S
SRAM Controller
LookUp Memory
Output Scheduler
O
Egress Buffer 3KB
G
424
422
Packet Processor A
Packet Processor B
P
Use Internal SRAM or SSRAM
Free Cell Manager
H
Header Queue 640 ~ 2560B
DRAM Controller
Packet Memory
D
F
1904
Ingress DMA
I
MAC Tx
A
MAC Rx
B
402
400

FIG. 19

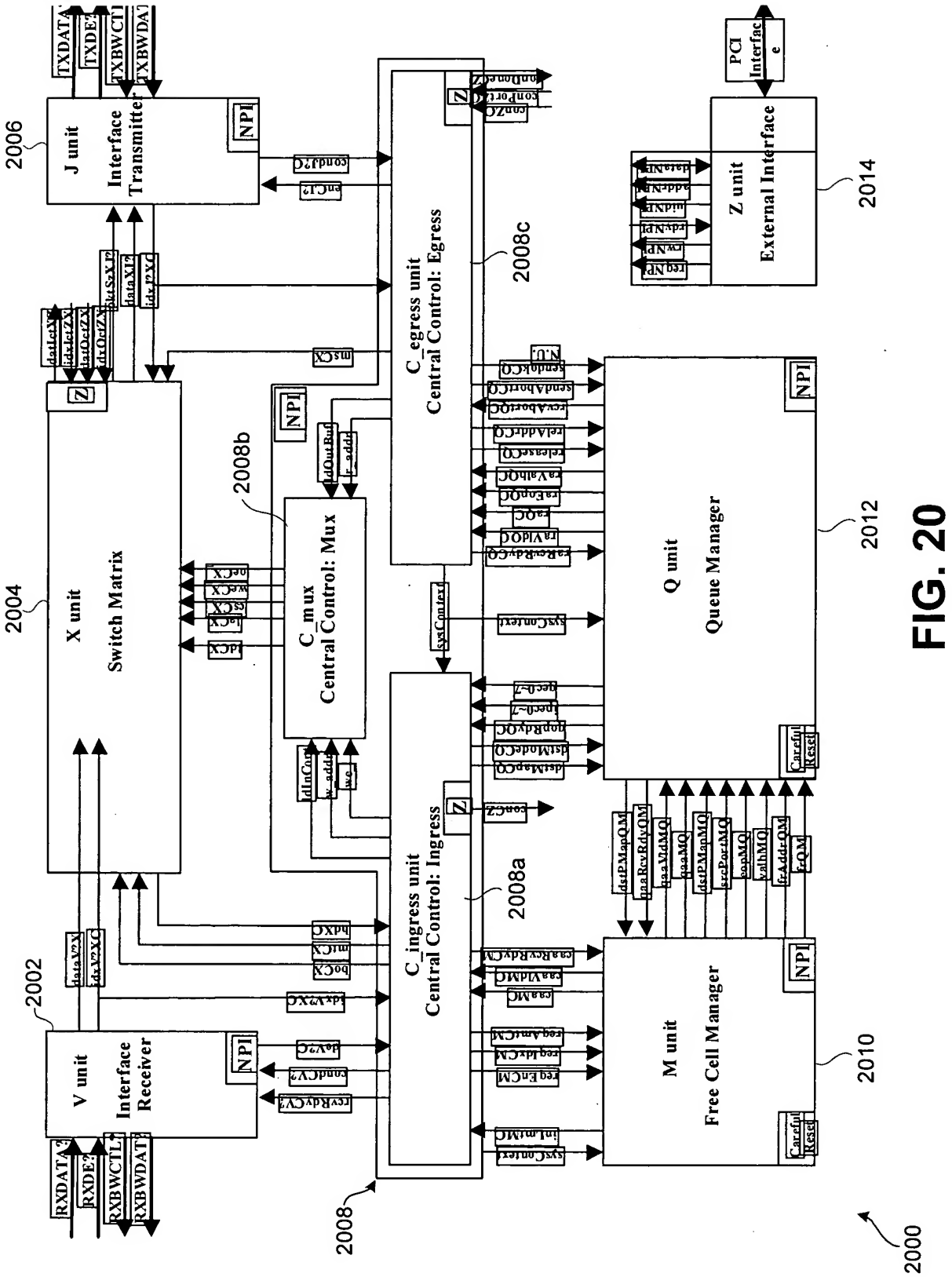
2000
2002
V unit
Interface
Receiver
2004
X unit
Switch Matrix
2006
J unit
Interface
Transmitter
2008
2008a
C_ingress unit
Central Control: Ingress
2008b
C_mux
Central Control: Mux
2008c
C_egress unit
Central Control: Egress
2010
M unit
Free Cell Manager
2012
Q unit
Queue Manager
2014
Z unit
External Interface
PCI
Interface
NPI
RXDATA
RXDEN
RXBWCTL
RXBWDATA
TXDATA
TXDEN
TXBWCTL
TXBWDATA

FIG. 20

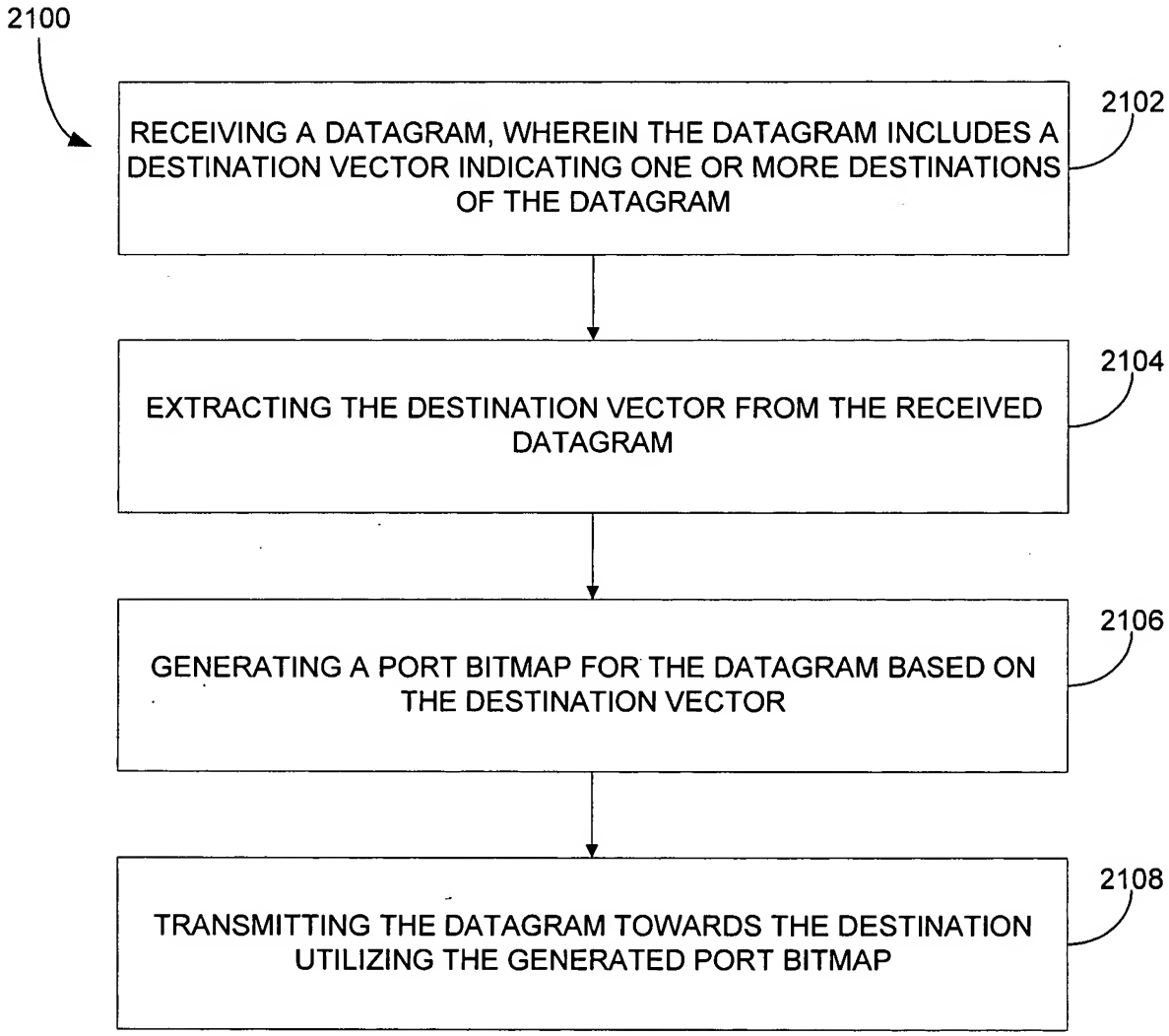
2100
RECEIVING A DATAGRAM, WHEREIN THE DATAGRAM INCLUDES A DESTINATION VECTOR INDICATING ONE OR MORE DESTINATIONS OF THE DATAGRAM
2102
EXTRACTING THE DESTINATION VECTOR FROM THE RECEIVED DATAGRAM
2104
GENERATING A PORT BITMAP FOR THE DATAGRAM BASED ON THE DESTINATION VECTOR
2106
TRANSMITTING THE DATAGRAM TOWARDS THE DESTINATION UTILIZING THE GENERATED PORT BITMAP
2108

FIG. 21

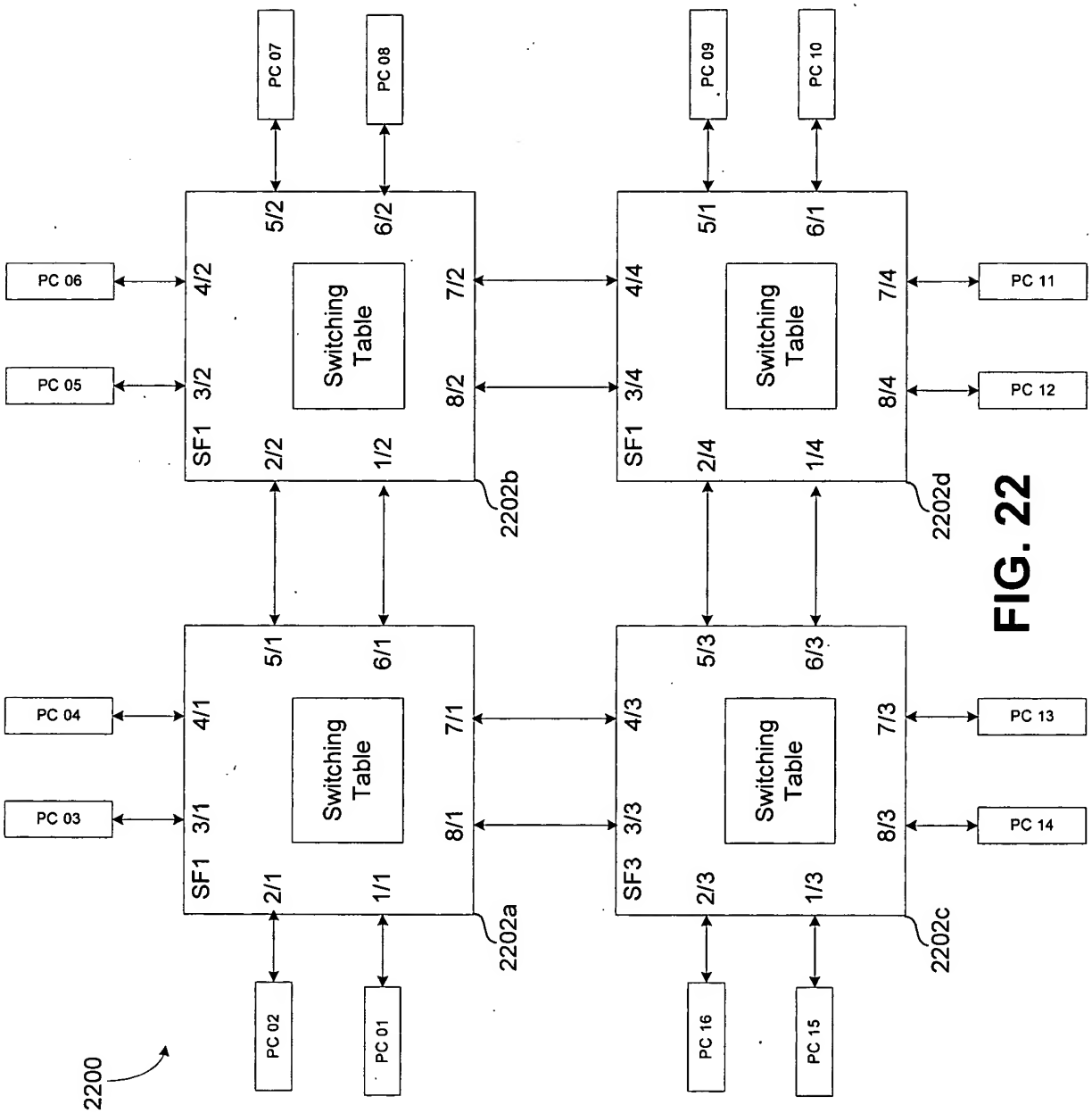
PC 07
PC 08
PC 09
PC 10
PC 06
PC 05
PC 11
PC 12
PC 04
PC 03
PC 13
PC 14
PC 02
PC 01
PC 16
PC 15
Switching Table
SF1
SF3
2202a
2202b
2202c
2202d
2200

FIG. 22

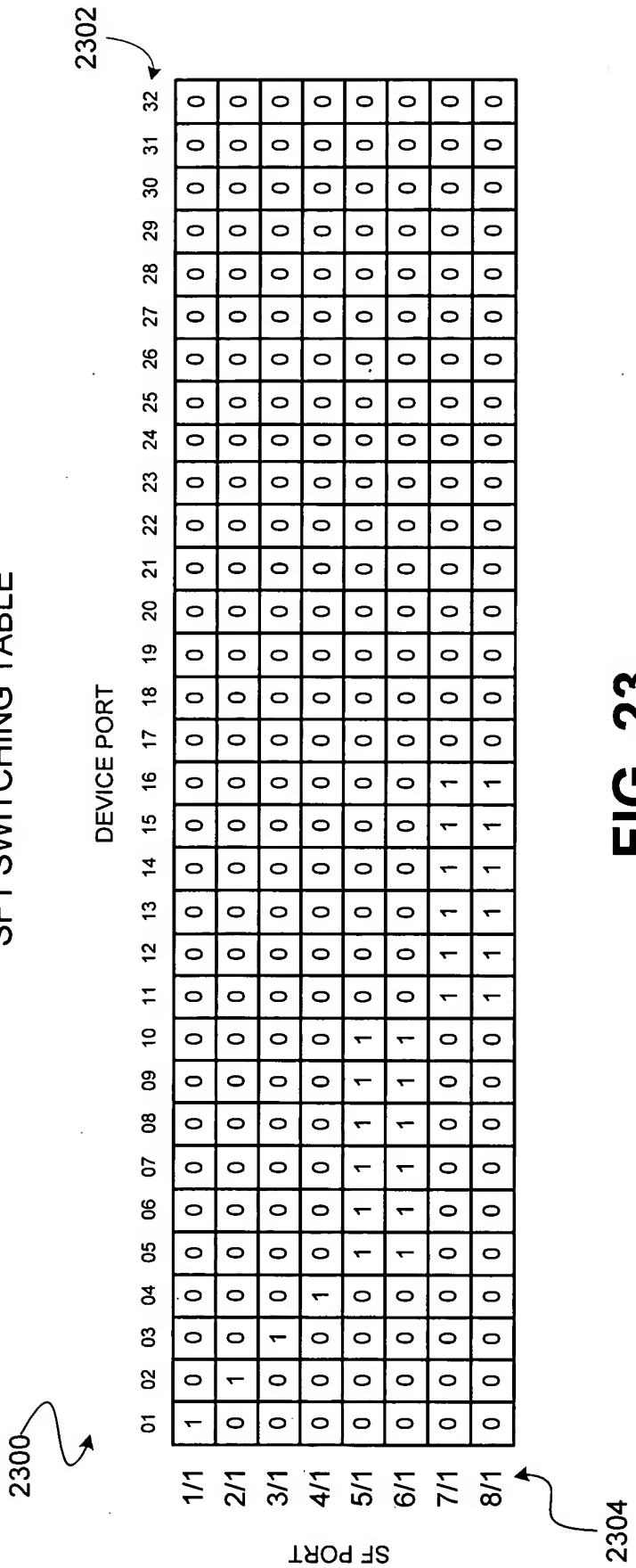
2302
SF1 SWITCHING TABLE
DEVICE PORT
01 02 03 04 05 06 07 08 09 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32
1/1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
2/1 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
3/1 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
4/1 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
5/1 0 0 0 0 1 1 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
6/1 0 0 0 0 1 1 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
7/1 0 0 0 0 0 0 0 0 0 0 1 1 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
8/1 0 0 0 0 0 0 0 0 0 0 1 1 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
SF PORT
2300
2304

FIG. 23

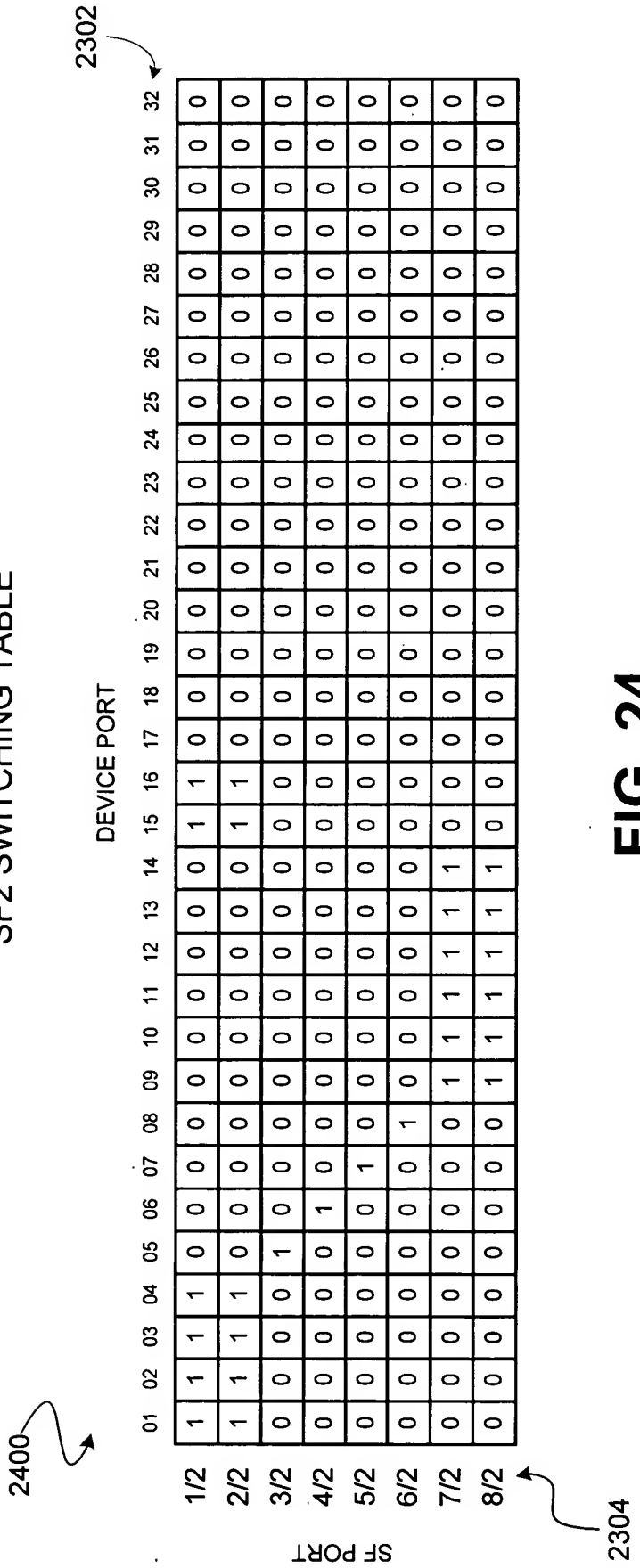
2302
SF2 SWITCHING TABLE
DEVICE PORT
01 02 03 04 05 06 07 08 09 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32
1/2 1 1 1 1 0 0 0 0 0 0 0 0 0 0 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
2/2 1 1 1 1 0 0 0 0 0 0 0 0 0 0 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
3/2 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
4/2 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
5/2 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
6/2 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
7/2 0 0 0 0 0 0 0 0 1 1 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
8/2 0 0 0 0 0 0 0 0 0 1 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
SF PORT
2400
2304

FIG. 24

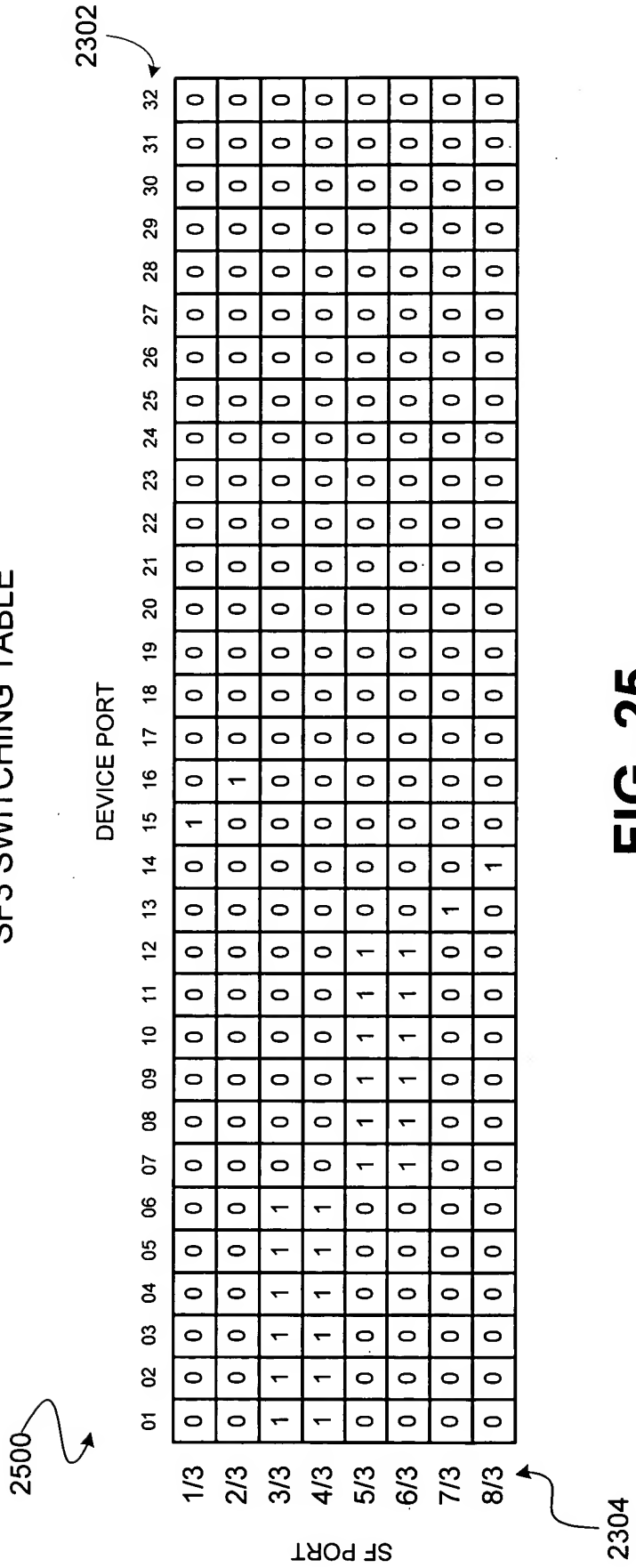
SF3 SWITCHING TABLE
DEVICE PORT
SF PORT
2302
2500
2304
01 02 03 04 05 06 07 08 09 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32
1/3 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
2/3 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
3/3 1 1 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
4/3 1 1 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
5/3 0 0 0 0 0 0 1 1 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
6/3 0 0 0 0 0 0 1 1 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
7/3 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
8/3 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

FIG. 25

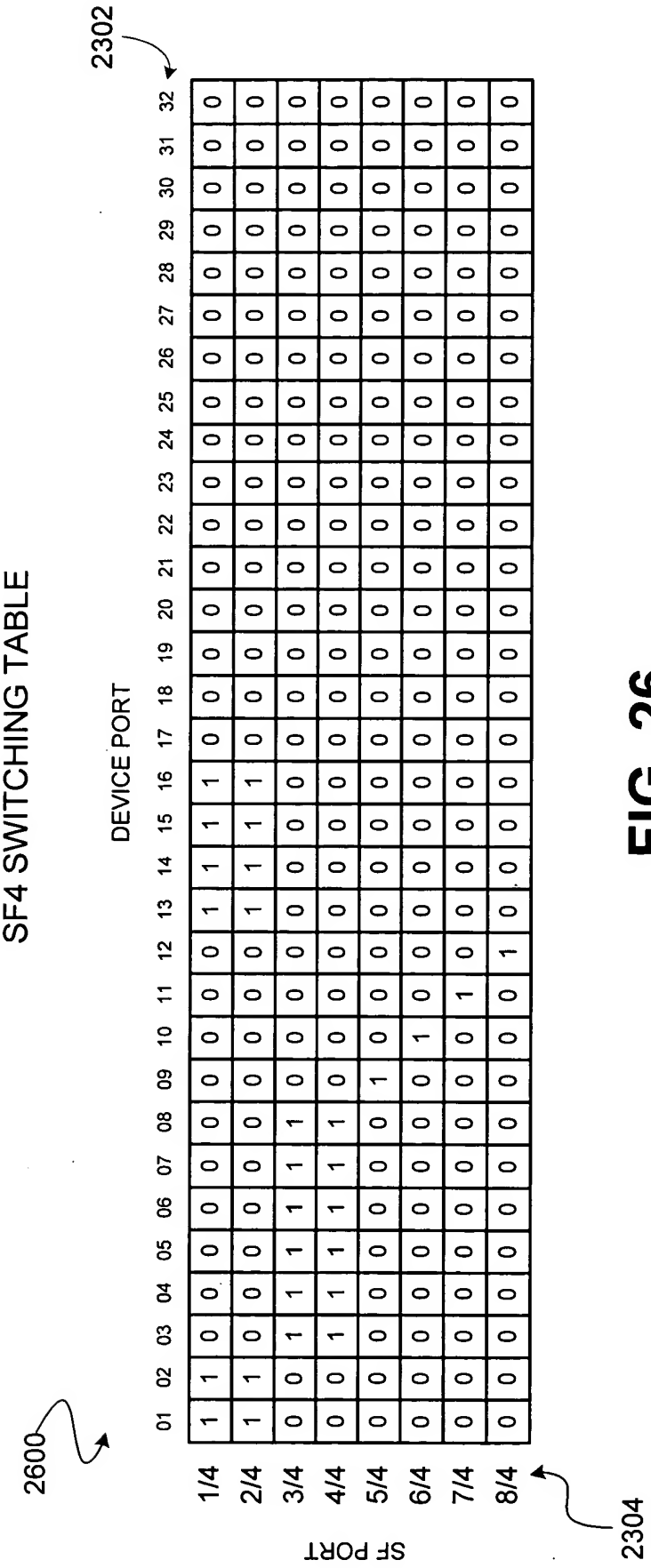
2302
SF4 SWITCHING TABLE
DEVICE PORT
01 02 03 04 05 06 07 08 09 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32
1/4 1 1 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
2/4 1 1 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
3/4 0 0 1 1 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
4/4 0 0 1 1 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
5/4 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
6/4 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
7/4 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
8/4 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
SF PORT
2600
2304

FIG. 26

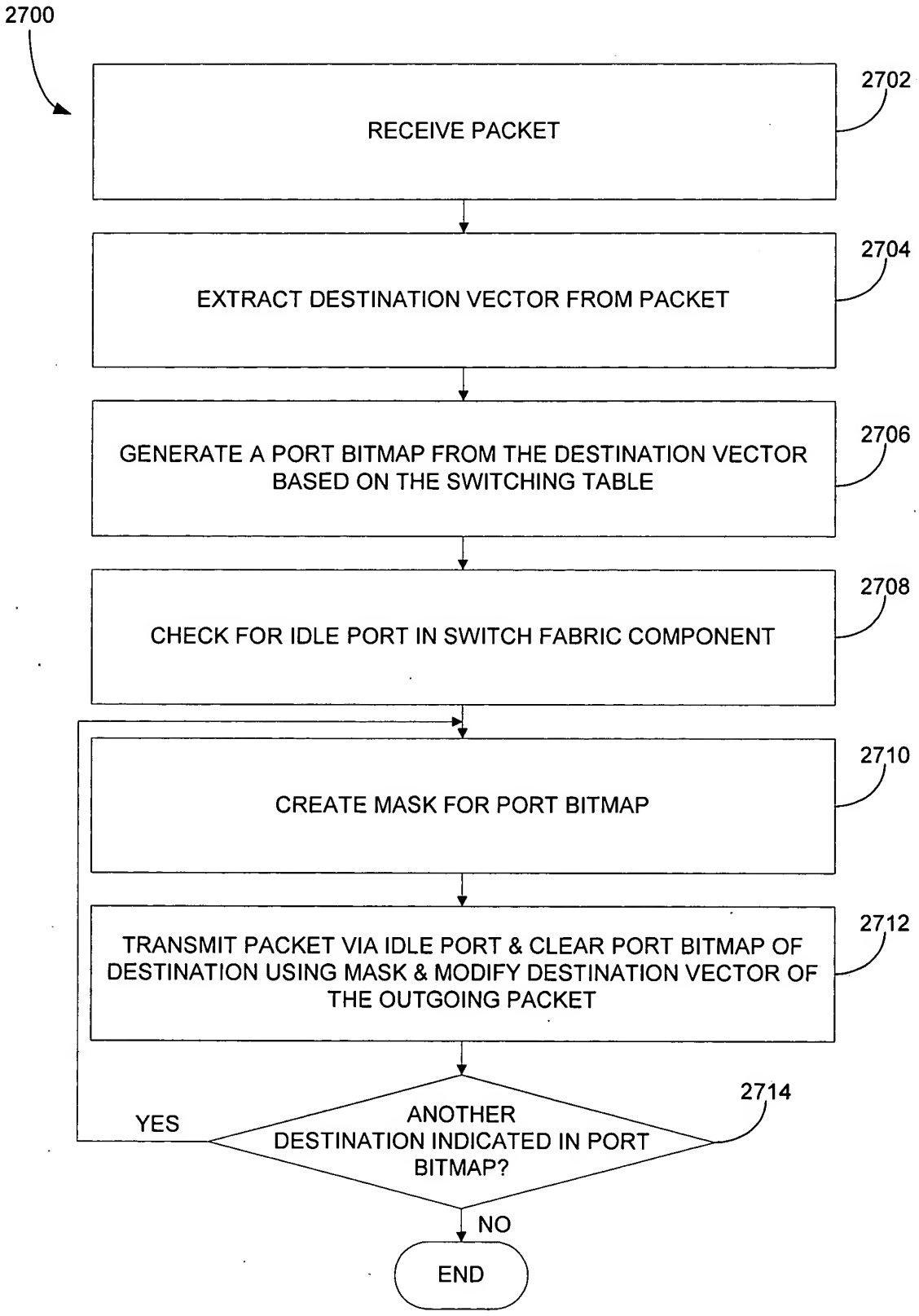
2700
RECEIVE PACKET
2702
EXTRACT DESTINATION VECTOR FROM PACKET
2704
GENERATE A PORT BITMAP FROM THE DESTINATION VECTOR BASED ON THE SWITCHING TABLE
2706
CHECK FOR IDLE PORT IN SWITCH FABRIC COMPONENT
2708
CREATE MASK FOR PORT BITMAP
2710
TRANSMIT PACKET VIA IDLE PORT & CLEAR PORT BITMAP OF DESTINATION USING MASK & MODIFY DESTINATION VECTOR OF THE OUTGOING PACKET
2712
ANOTHER DESTINATION INDICATED IN PORT BITMAP?
2714
YES
NO
END

FIG. 27

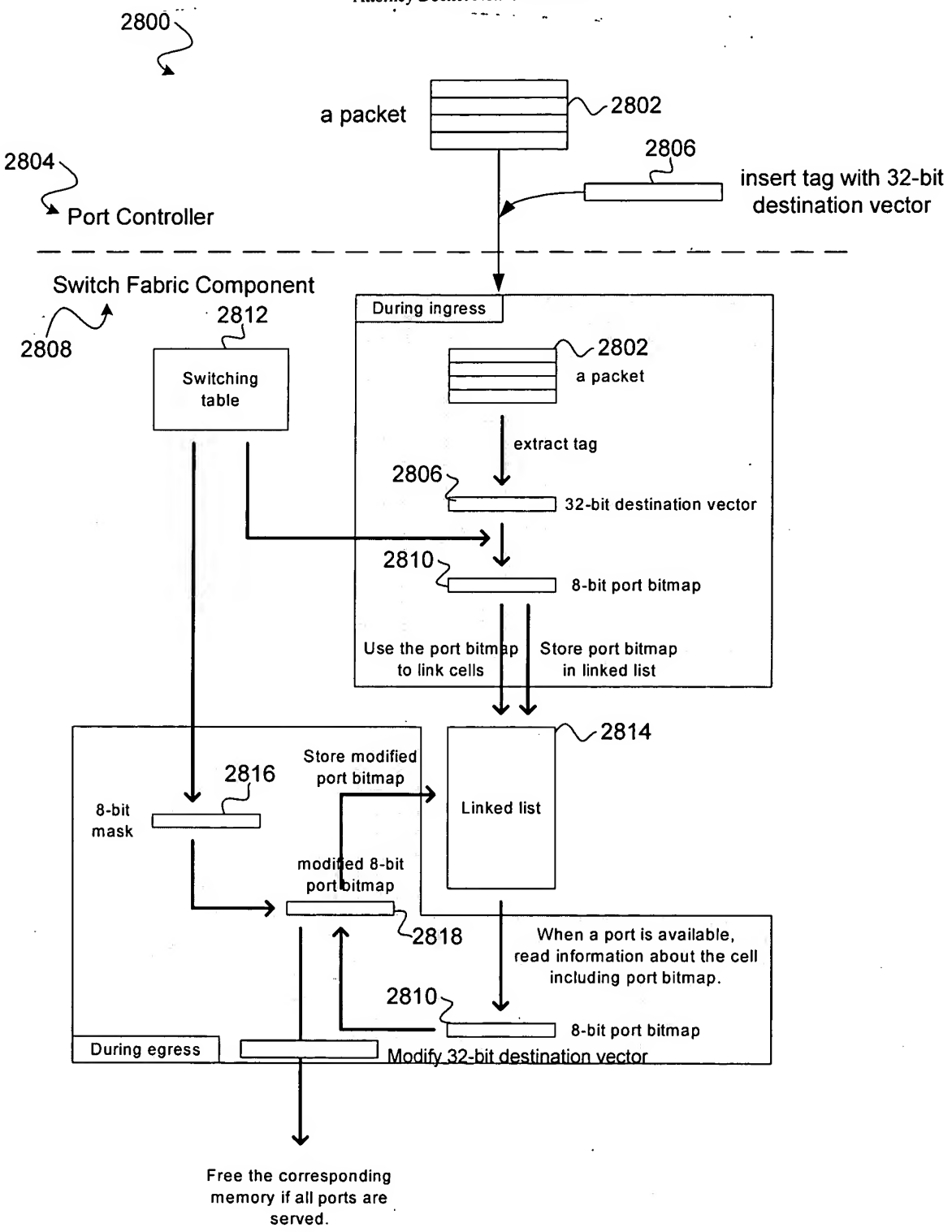
2800
a packet
2802
2804
2806
insert tag with 32-bit destination vector
Port Controller
Switch Fabric Component
2808
2812
Switching table
During ingress
2802
a packet
extract tag
2806
32-bit destination vector
2810
8-bit port bitmap
Use the port bitmap to link cells
Store port bitmap in linked list
2814
Linked list
Store modified port bitmap
2816
8-bit mask
modified 8-bit port bitmap
2818
When a port is available, read information about the cell including port bitmap.
2810
8-bit port bitmap
During egress
Modify 32-bit destination vector
Free the corresponding memory if all ports are served.

FIG. 28

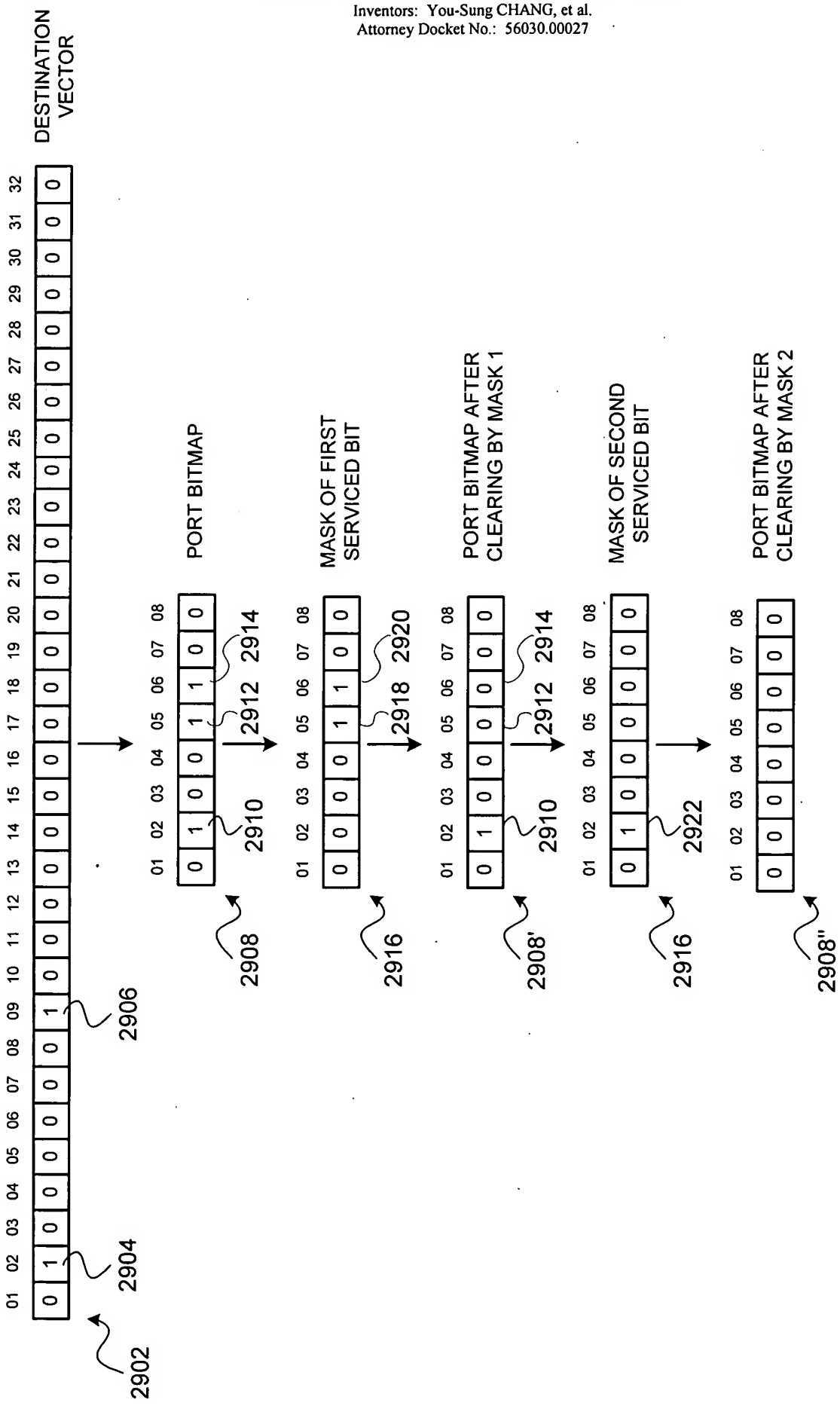
DESTINATION VECTOR
01 02 03 04 05 06 07 08 09 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32
0 1 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
2902
2904
2906
PORT BITMAP
01 02 03 04 05 06 07 08
0 1 0 0 1 1 0 0
2908
2910
2912
2914
MASK OF FIRST SERVICED BIT
01 02 03 04 05 06 07 08
0 0 0 0 1 1 0 0
2916
2918
2920
PORT BITMAP AFTER CLEARING BY MASK 1
01 02 03 04 05 06 07 08
0 1 0 0 0 0 0 0
2908'
2910
2912
2914
MASK OF SECOND SERVICED BIT
01 02 03 04 05 06 07 08
0 1 0 0 0 0 0 0
2916
2922
PORT BITMAP AFTER CLEARING BY MASK 2
01 02 03 04 05 06 07 08
0 0 0 0 0 0 0 0
2908''

FIG. 29

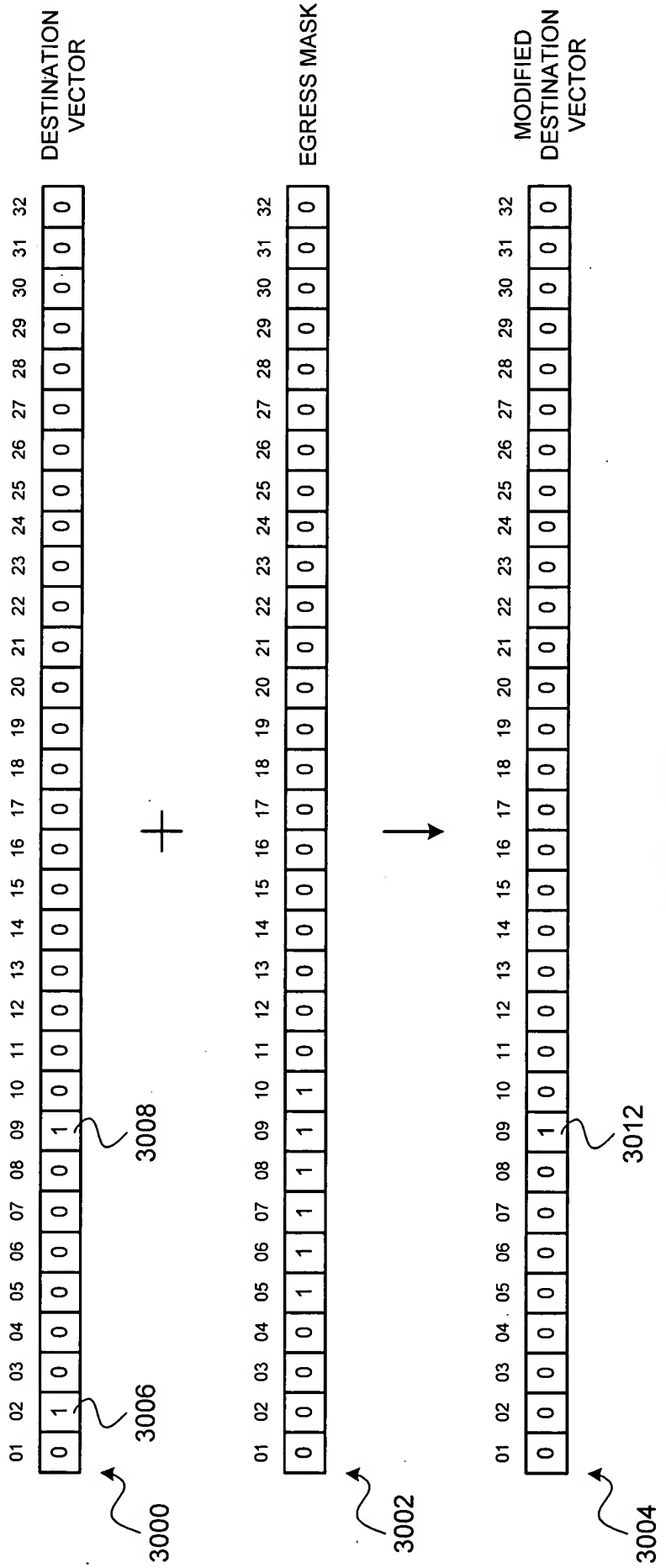
DESTINATION VECTOR
EGRESS MASK
MODIFIED DESTINATION VECTOR
01 02 03 04 05 06 07 08 09 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32
0 1 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 1 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
3000
3002
3004
3006
3008
3012

FIG. 30

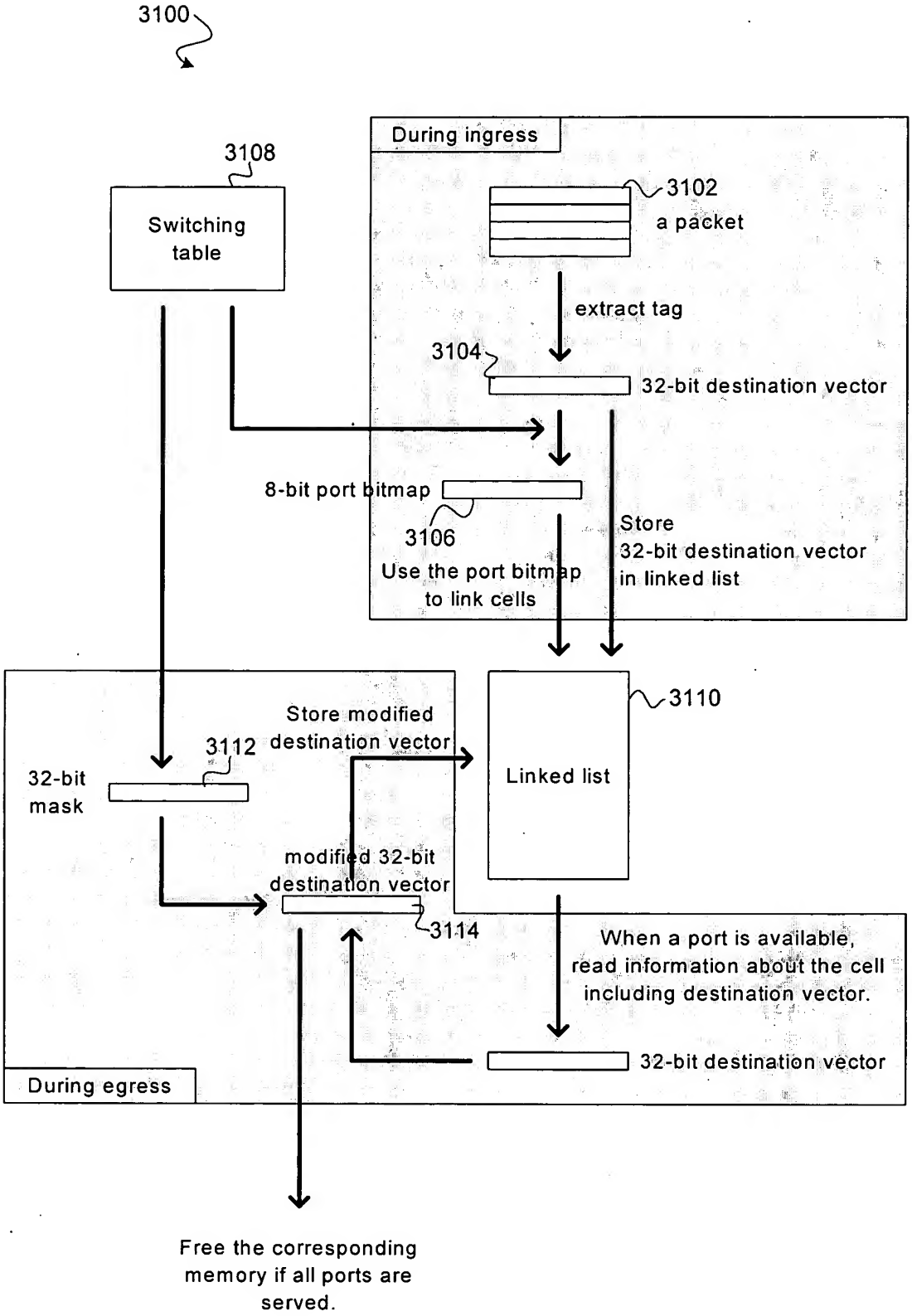
3100
During ingress
3108
Switching table
3102
a packet
extract tag
3104
32-bit destination vector
8-bit port bitmap
3106
Use the port bitmap to link cells
Store 32-bit destination vector in linked list
3110
Linked list
Store modified destination vector
3112
32-bit mask
modified 32-bit destination vector
3114
When a port is available, read information about the cell including destination vector.
32-bit destination vector
During egress
Free the corresponding memory if all ports are served.

FIG. 31

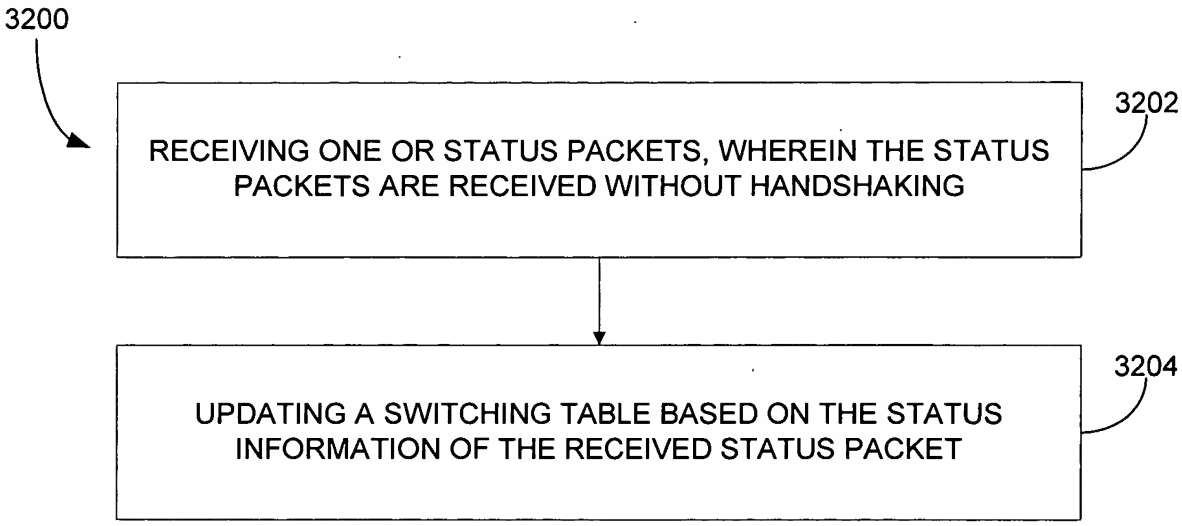
3200
RECEIVING ONE OR STATUS PACKETS, WHEREIN THE STATUS PACKETS ARE RECEIVED WITHOUT HANDSHAKING
3202
UPDATING A SWITCHING TABLE BASED ON THE STATUS INFORMATION OF THE RECEIVED STATUS PACKET
3204

FIG. 32

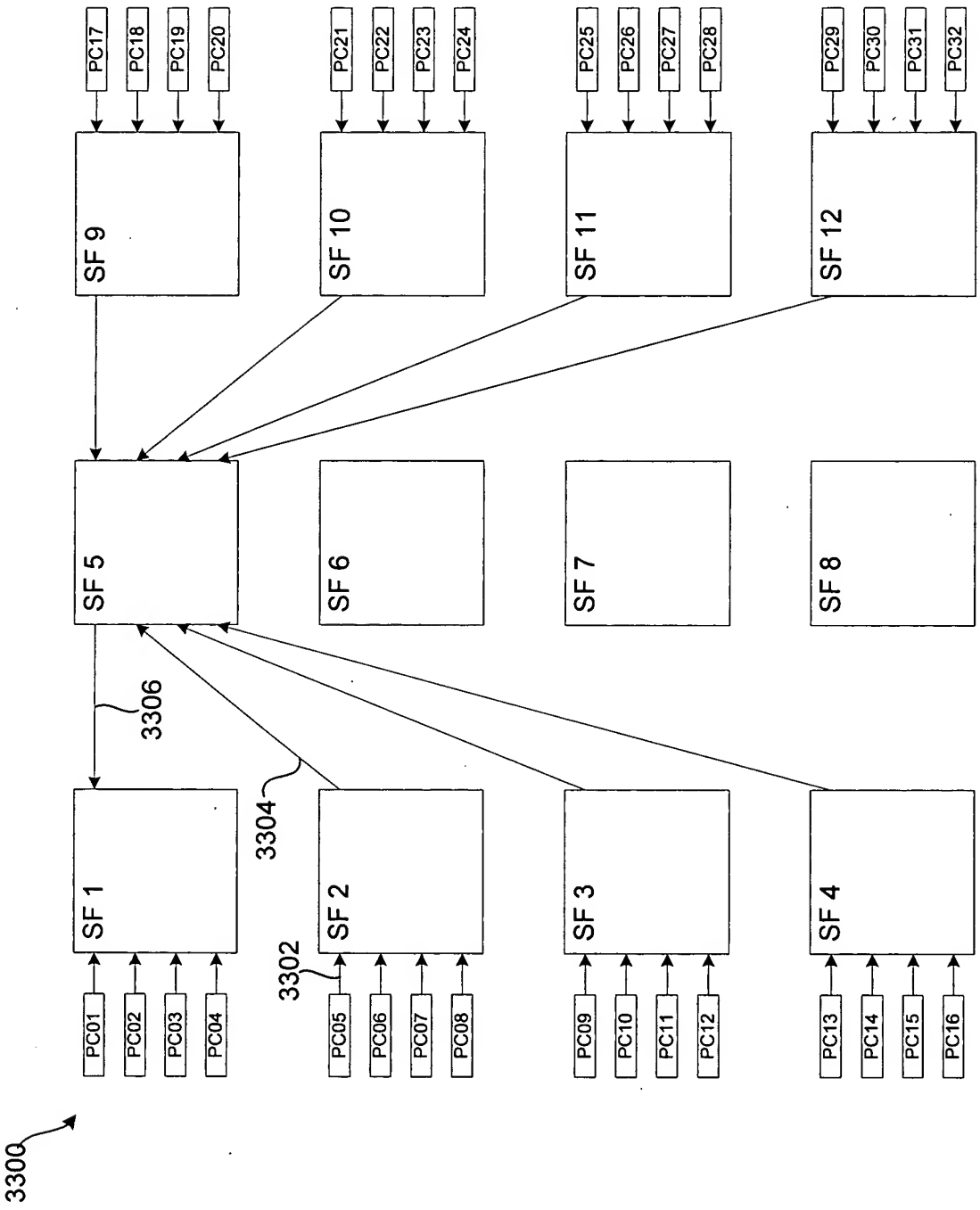
PC17
PC18
PC19
PC20
PC21
PC22
PC23
PC24
PC25
PC26
PC27
PC28
PC29
PC30
PC31
PC32
SF 9
SF 10
SF 11
SF 12
SF 5
SF 6
SF 7
SF 8
3306
3304
3302
SF 1
SF 2
SF 3
SF 4
PC01
PC02
PC03
PC04
PC05
PC06
PC07
PC08
PC09
PC10
PC11
PC12
PC13
PC14
PC15
PC16
3300

FIG. 33

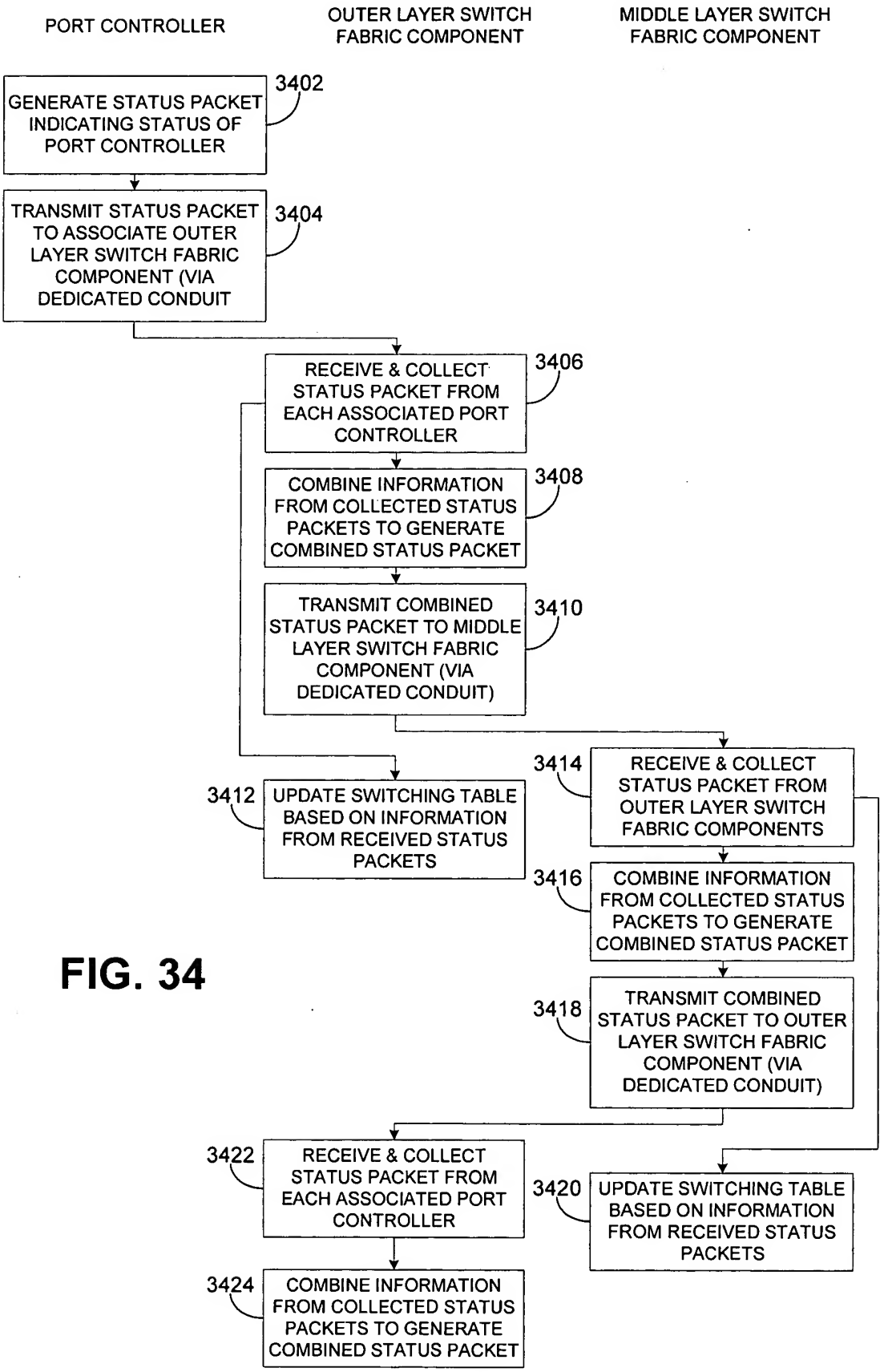
PORT CONTROLLER
OUTER LAYER SWITCH FABRIC COMPONENT
MIDDLE LAYER SWITCH FABRIC COMPONENT
GENERATE STATUS PACKET INDICATING STATUS OF PORT CONTROLLER
3402
TRANSMIT STATUS PACKET TO ASSOCIATE OUTER LAYER SWITCH FABRIC COMPONENT (VIA DEDICATED CONDUIT
3404
RECEIVE & COLLECT STATUS PACKET FROM EACH ASSOCIATED PORT CONTROLLER
3406
COMBINE INFORMATION FROM COLLECTED STATUS PACKETS TO GENERATE COMBINED STATUS PACKET
3408
TRANSMIT COMBINED STATUS PACKET TO MIDDLE LAYER SWITCH FABRIC COMPONENT (VIA DEDICATED CONDUIT)
3410
3412
UPDATE SWITCHING TABLE BASED ON INFORMATION FROM RECEIVED STATUS PACKETS
3414
RECEIVE & COLLECT STATUS PACKET FROM OUTER LAYER SWITCH FABRIC COMPONENTS
3416
COMBINE INFORMATION FROM COLLECTED STATUS PACKETS TO GENERATE COMBINED STATUS PACKET
3418
TRANSMIT COMBINED STATUS PACKET TO OUTER LAYER SWITCH FABRIC COMPONENT (VIA DEDICATED CONDUIT)
3422
RECEIVE & COLLECT STATUS PACKET FROM EACH ASSOCIATED PORT CONTROLLER
3420
UPDATE SWITCHING TABLE BASED ON INFORMATION FROM RECEIVED STATUS PACKETS
3424
COMBINE INFORMATION FROM COLLECTED STATUS PACKETS TO GENERATE COMBINED STATUS PACKET

FIG. 34

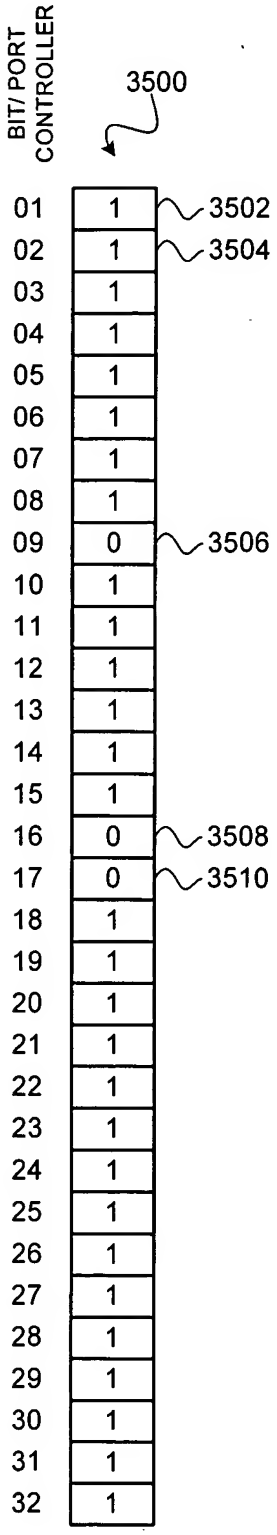
BIT/ PORT CONTROLLER
3500
01 1
02 1
03 1
04 1
05 1
06 1
07 1
08 1
09 0
10 1
11 1
12 1
13 1
14 1
15 1
16 0
17 0
18 1
19 1
20 1
21 1
22 1
23 1
24 1
25 1
26 1
27 1
28 1
29 1
30 1
31 1
32 1
3502
3504
3506
3508
3510

FIG. 35

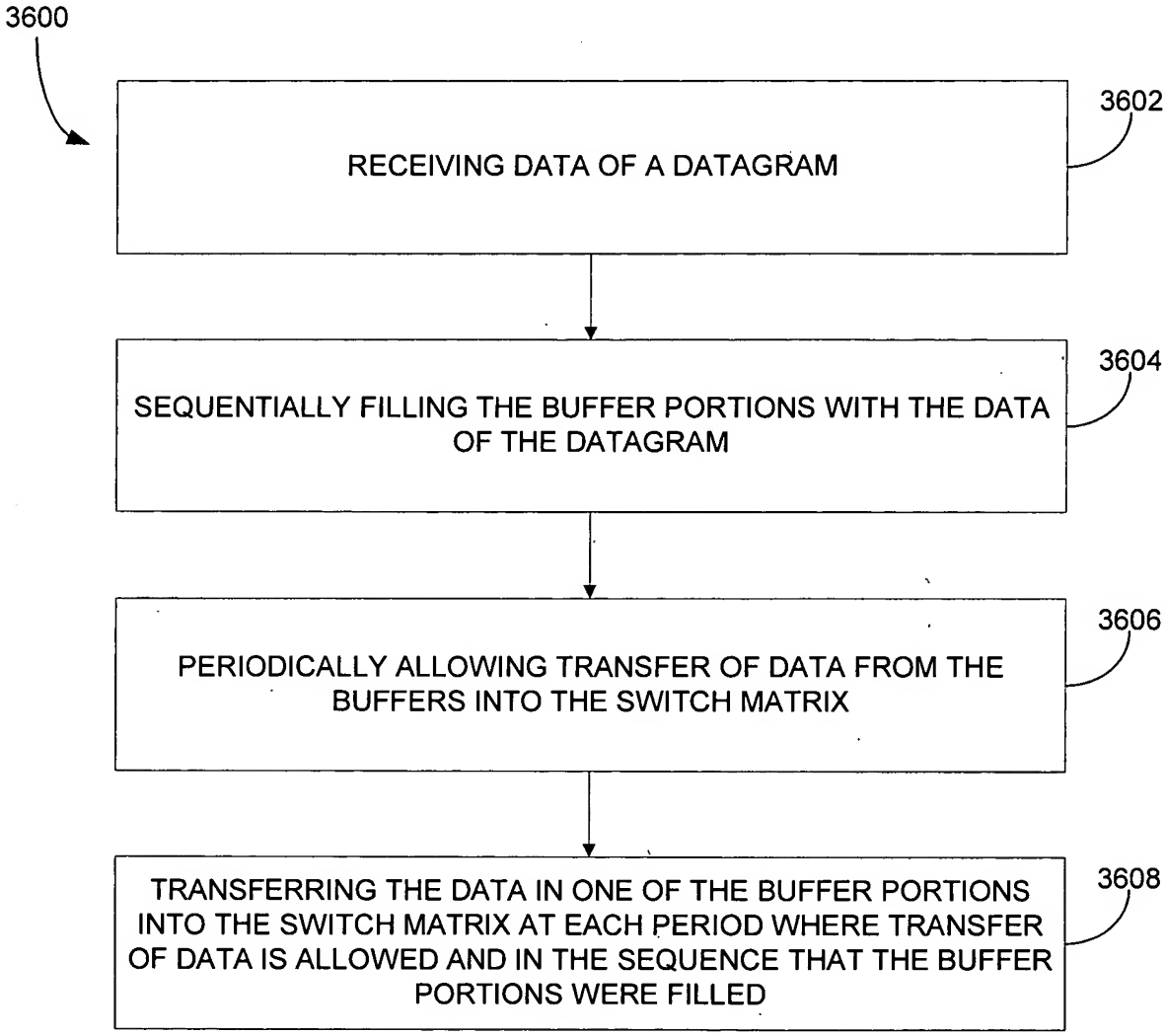
3600
RECEIVING DATA OF A DATAGRAM
3602
SEQUENTIALLY FILLING THE BUFFER PORTIONS WITH THE DATA OF THE DATAGRAM
3604
PERIODICALLY ALLOWING TRANSFER OF DATA FROM THE BUFFERS INTO THE SWITCH MATRIX
3606
TRANSFERRING THE DATA IN ONE OF THE BUFFER PORTIONS INTO THE SWITCH MATRIX AT EACH PERIOD WHERE TRANSFER OF DATA IS ALLOWED AND IN THE SEQUENCE THAT THE BUFFER PORTIONS WERE FILLED
3608

FIG. 36

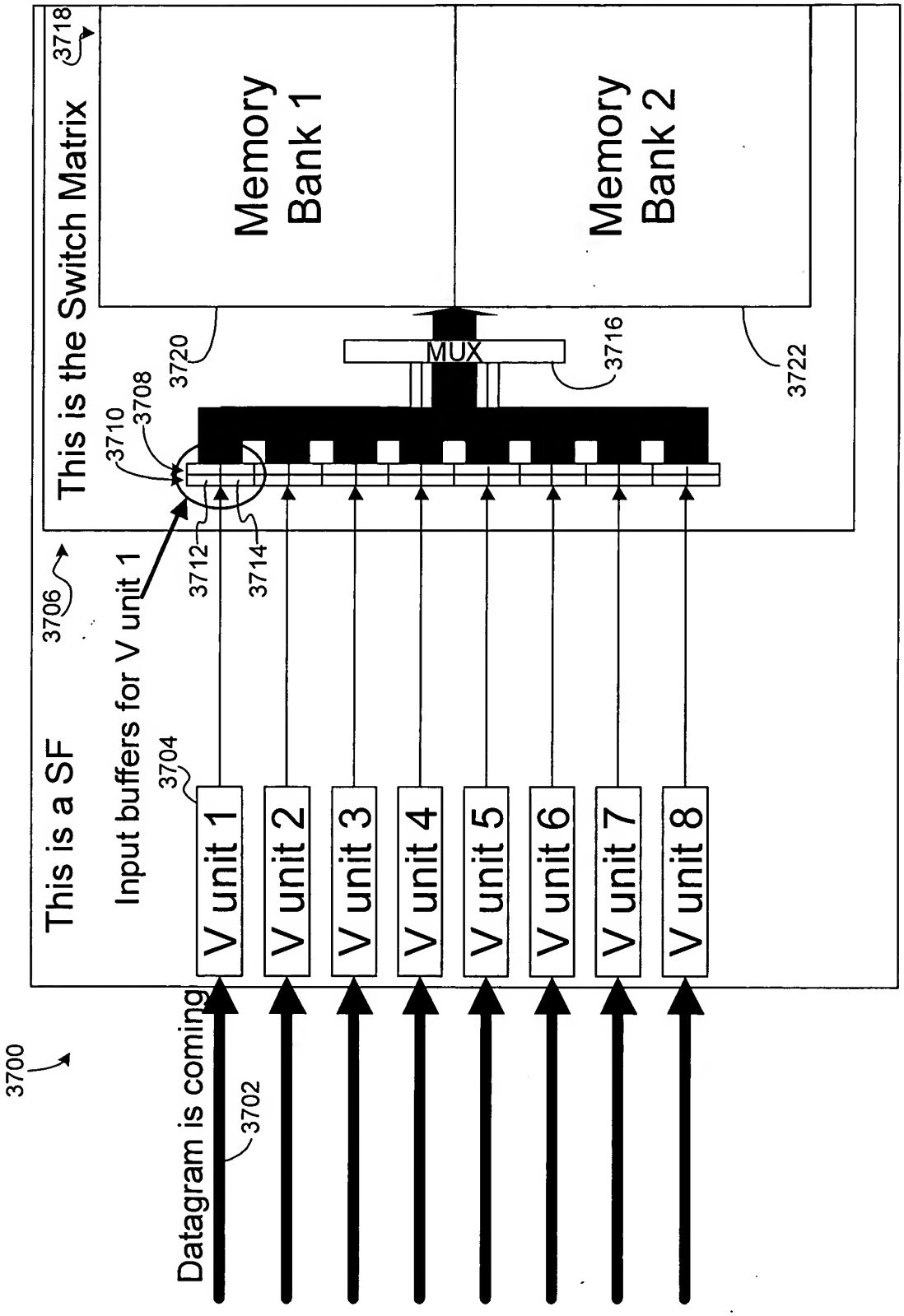
3700
3702
Datagram is coming
This is a SF
3704
V unit 1
V unit 2
V unit 3
V unit 4
V unit 5
V unit 6
V unit 7
V unit 8
3706
Input buffers for V unit 1
3712
3714
This is the Switch Matrix
3718
Memory Bank 1
Memory Bank 2
3720
3722
3708
3710
MUX
3716

FIG. 37

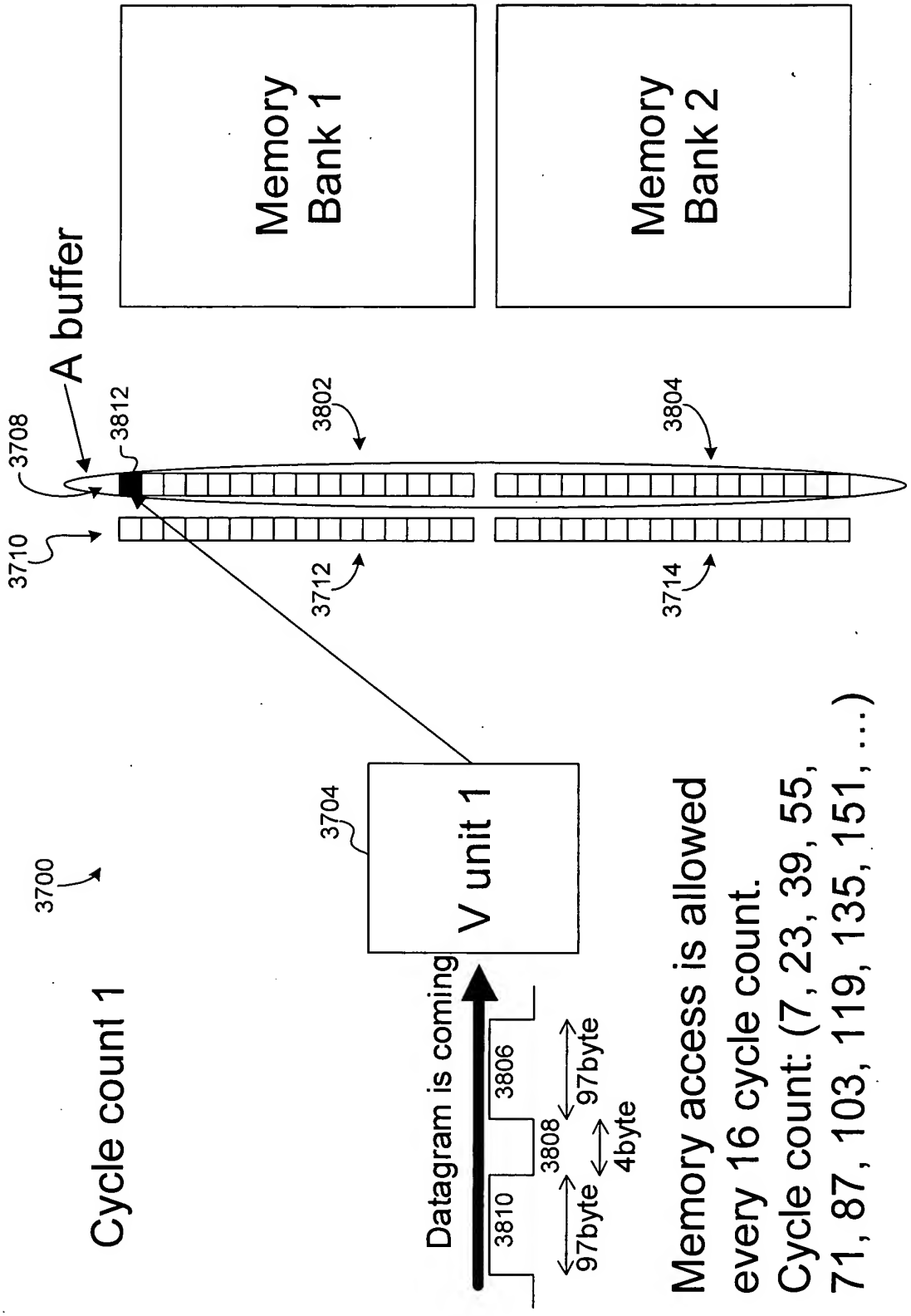
Memory Bank 1
Memory Bank 2
A buffer
3708
3812
3802
3804
3710
3712
3714
3704
V unit 1
3700
Cycle count 1
Datagram is coming
3806
3808
3810
97byte
4byte
97byte
Memory access is allowed every 16 cycle count.
Cycle count: (7, 23, 39, 55, 71, 87, 103, 119, 135, 151, ...)

FIG. 38

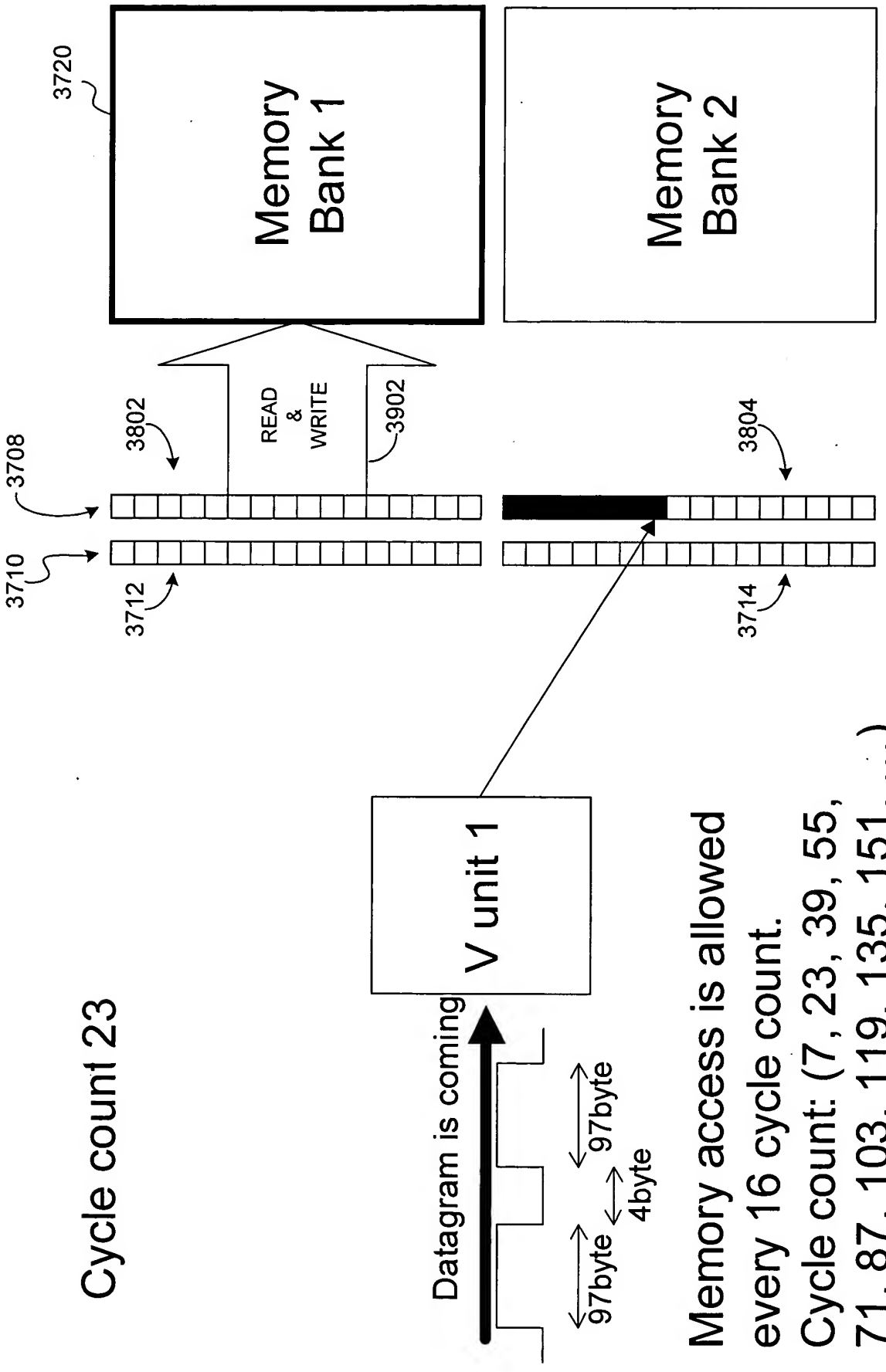
3720
Memory Bank 1
Memory Bank 2
3708
3802
READ & WRITE
3902
3804
3710
3712
3714
V unit 1
Datagram is coming
97byte
4byte
97byte
Cycle count 23
Memory access is allowed every 16 cycle count.
Cycle count: (7, 23, 39, 55, 71, 87, 103, 119, 135, 151, ...)

FIG. 39

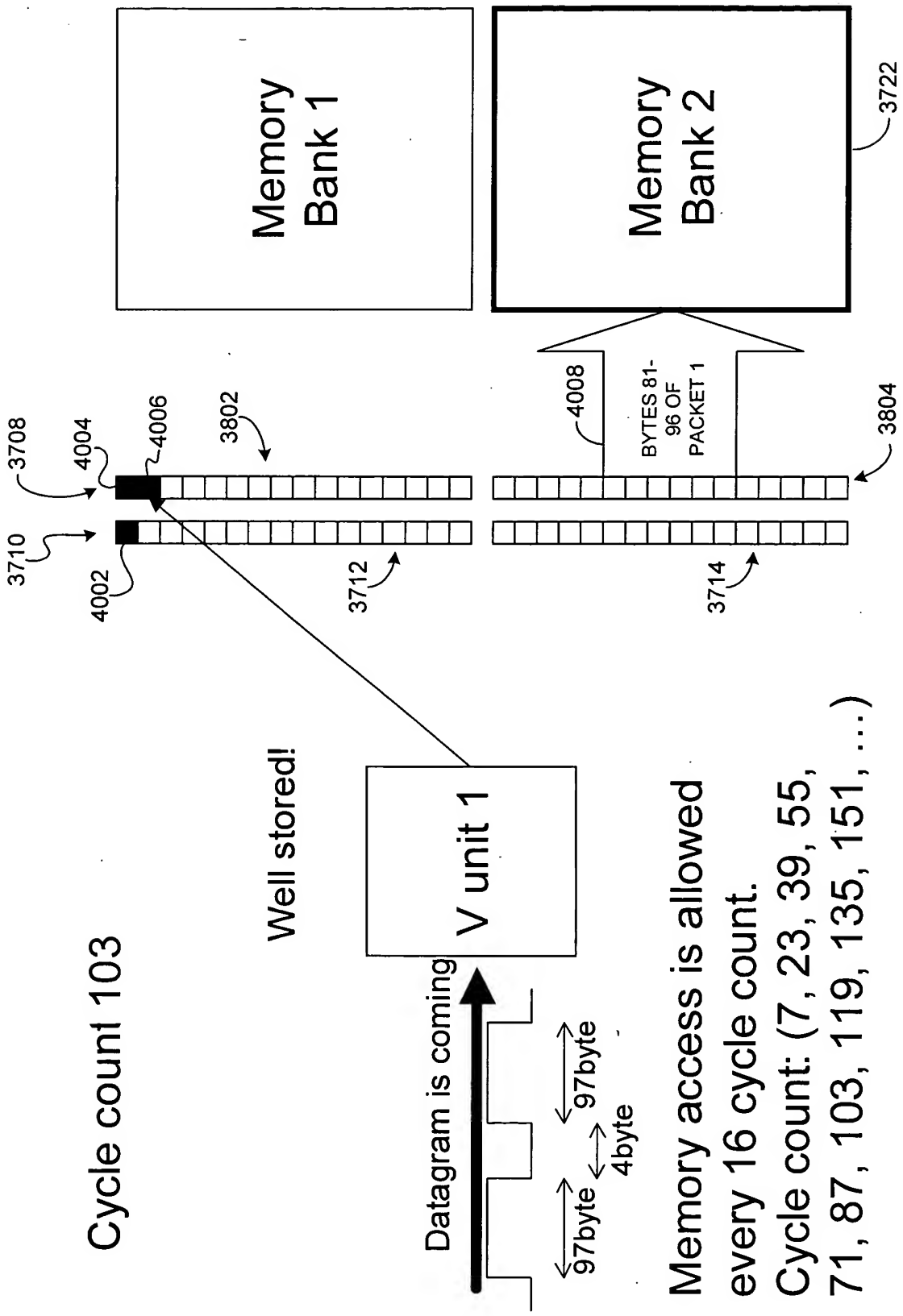
Memory Bank 1
Memory Bank 2
3722
3708
4004
4006
3802
4008
BYTES 81-96 OF PACKET 1
3804
3710
4002
3712
3714
Cycle count 103
Well stored!
V unit 1
Datagram is coming
97byte
4byte
97byte
Memory access is allowed every 16 cycle count.
Cycle count: (7, 23, 39, 55, 71, 87, 103, 119, 135, 151, ...)

FIG. 40

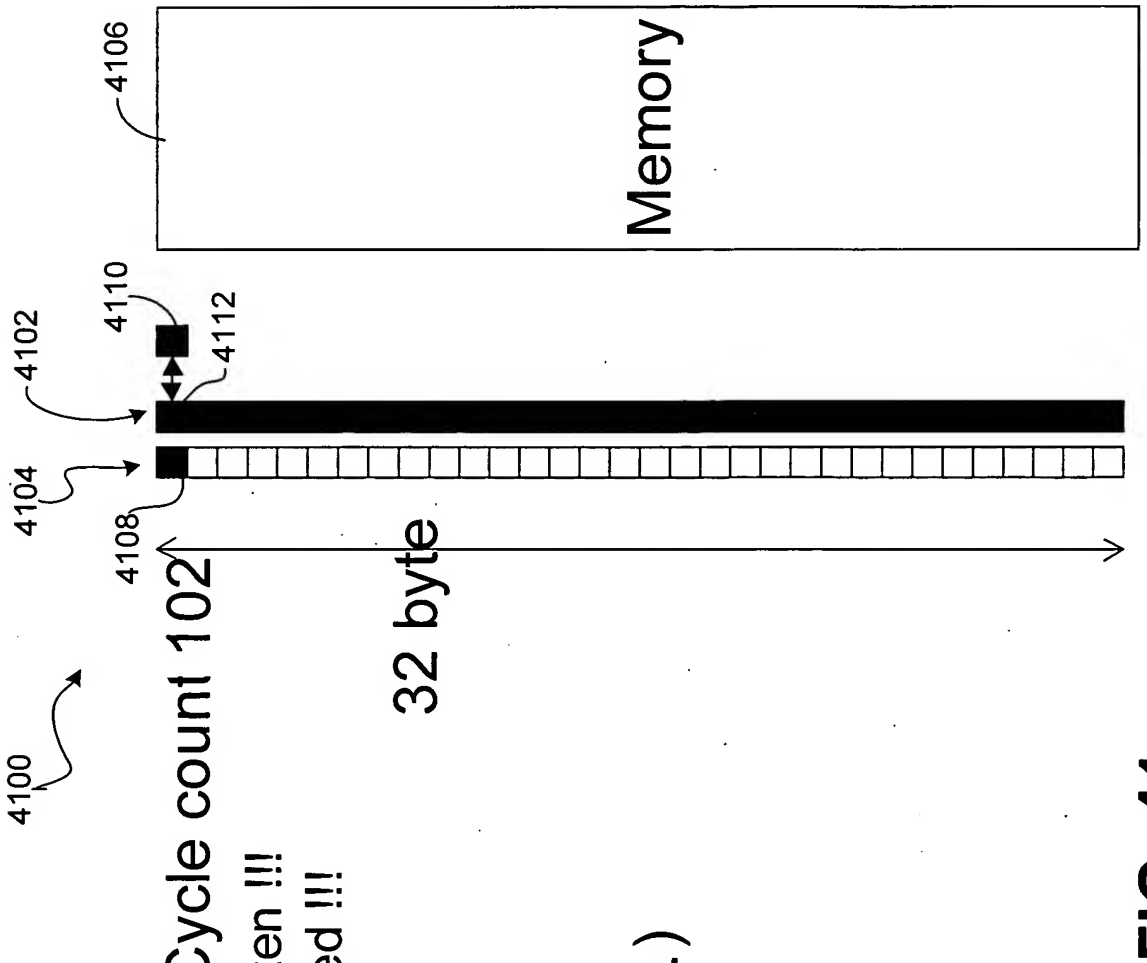
4106
Memory
4110
4112
4102
4104
4108
4100
Cycle count 102
32 byte
Overwritten !!!
Damaged !!!
Memory access is allowed
every 16 cycle count.
Cycle count: (7, 23, 39, 55,
71, 87, 103, 119, 135, 151, ...)
Datagram is coming
97byte
4byte
97byte

FIG. 41

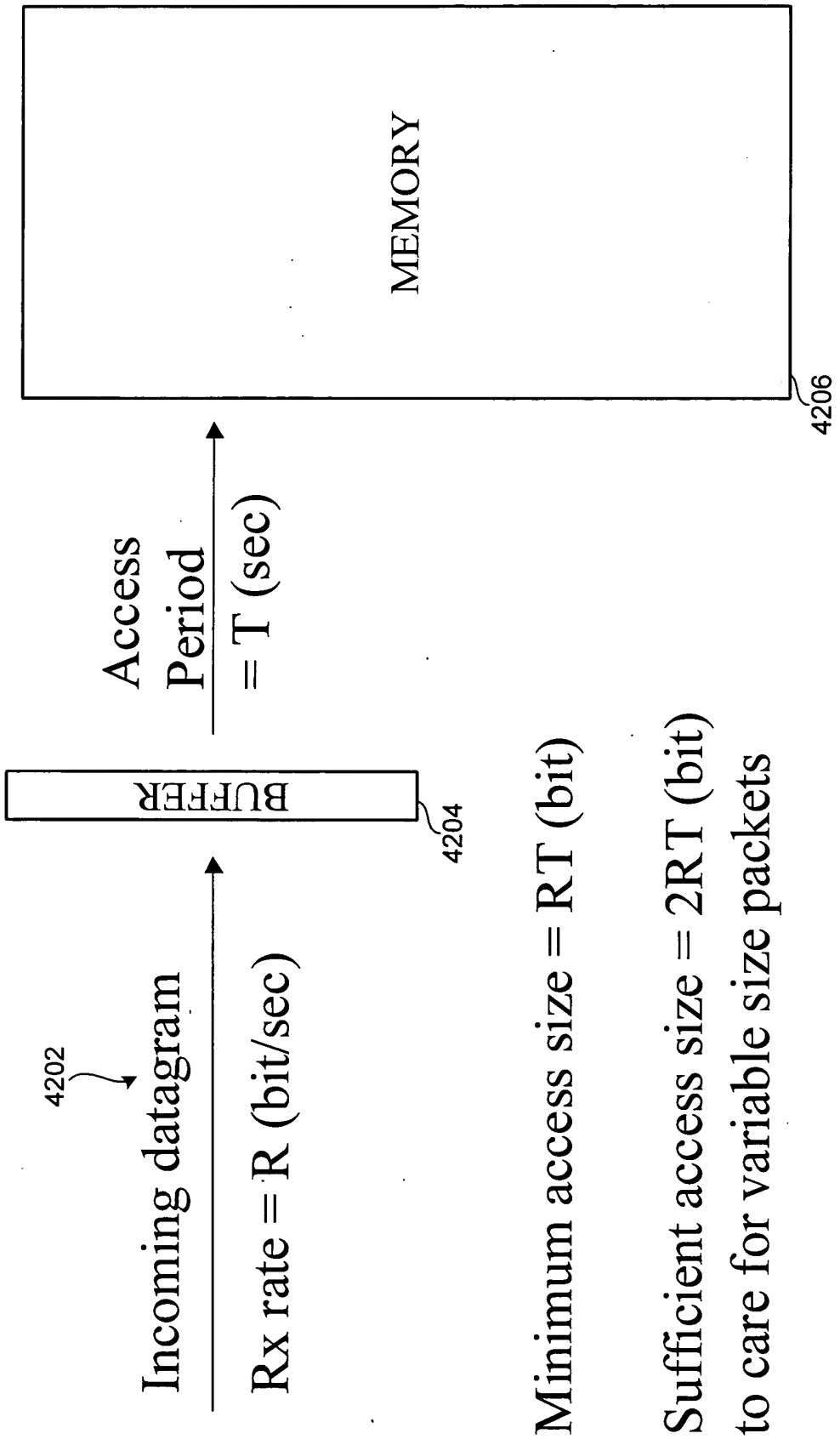
Incoming datagram
4202
Rx rate = R (bit/sec)
BUFFER
4204
Access Period = T (sec)
MEMORY
4206
Minimum access size = RT (bit)
Sufficient access size = 2RT (bit)
to care for variable size packets

FIG. 42

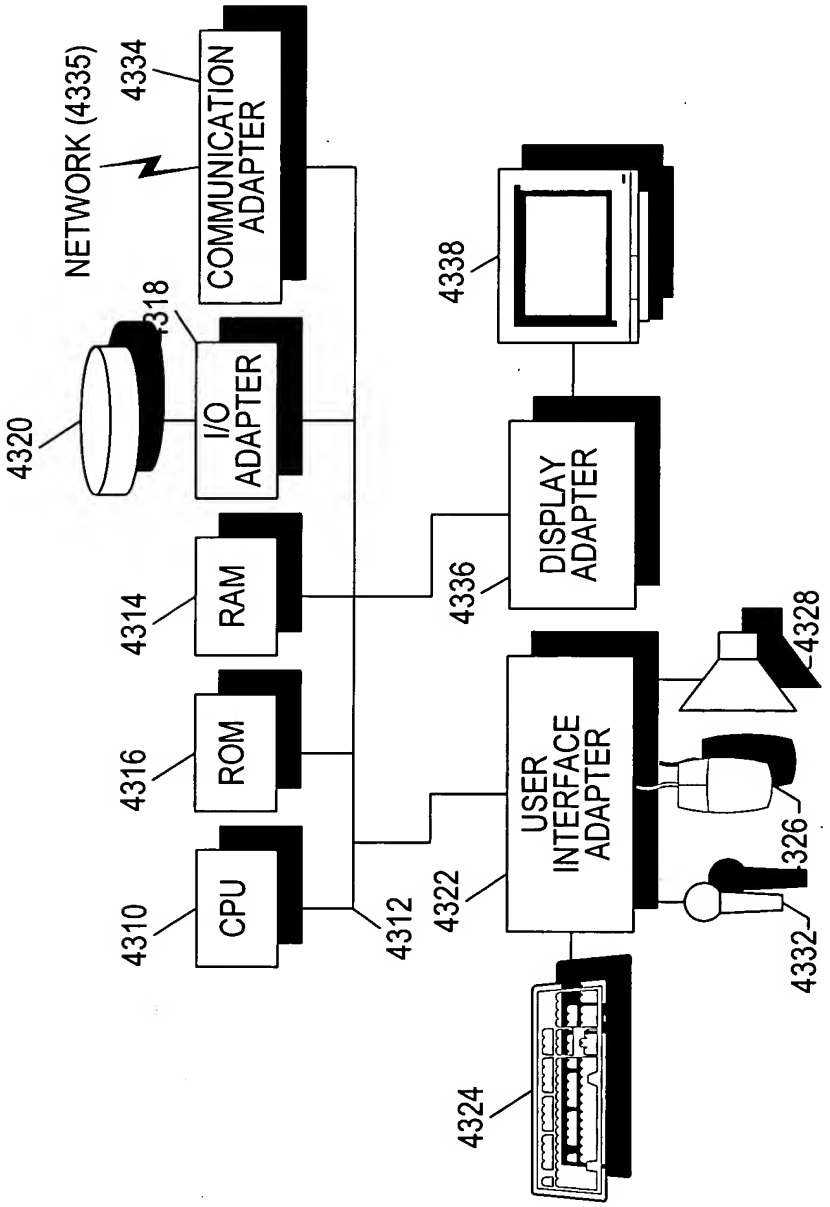
NETWORK (4335)
4334
COMMUNICATION ADAPTER
4318
I/O ADAPTER
4320
4314
RAM
4316
ROM
4310
CPU
4312
4338
DISPLAY ADAPTER
4336
4328
USER INTERFACE ADAPTER
4322
326
4332
4324

FIG. 43